Head Start and Mothers' Work: Free Child Care or Something More?

by

Ariel Marek Pihl
University of Gothenburg

CES 18-13 **March, 2018**

Abstract

Head Start is the largest public pre-school program in the US, but it provides many additional services to families. This paper uses a discontinuity in grant writing assistance in the first year of the Head Start program to identify impacts on the work and welfare usage of mothers. Using restricted Decennial Census and administrative AFDC data I find that Head Start decreases employment rates and hours worked per week for single mothers. I also find a suggestive increase in welfare receipt for single mothers which is confirmed by an increase in the share of administrative welfare case-files that are single mother households. For all mothers combined there are no significant changes in work or welfare use. I also estimate long-run impacts, 10 years after a woman's child was eligible for Head Start. I find large and persistent declines in work for both non-white mothers and single mothers, accompanied by an increase in public assistance income and return to school. I argue that this is consistent with the 1960's era Head Start program's focus on encouraging quality parenting, parent participation and helping families access all benefits for which they were eligible.

Keyword*:*

JEL Classification: H53, I32, I38, J22

[*] I am grateful to Doug Miller, Marianne Bitler and Marianne Page for their guidance, support and insightful comments. I also thank Martha Bailey for supporting my work with the restricted Census and Andrew Goodman-Bacon for sharing the AFDC data. I thank Scott Carrell, Eric Muehlegger, Cassandra Hart, Giovanni Peri, Michel Grosz, Esra Kose and the participants in the UC Davis Brown Bag and All-Cal Labor Economics Conference for their feedback. All errors are my own.

DISCLAIMER: Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

*Department of Economics, University of Gothenburg, E-mail: ariel.pihl@gu.se.

1 Introduction

Head Start, the largest preschool education program in America, has served the country's poor children for more than 50 years. Today, federal spending on Head Start totals $8.6 billion per year, accounting for the majority of federal spending on early childhood education, and surpassing total state and local spending on such services (Department of Health and Human Services, 2015; Barnett et al., 2016).[1] Since its early years, this comprehensive preschool program has had dedicated supporters and critics. Justifiably, the research into Head Start's impacts has focused on the children it serves. However, Head Start has always been a program with broad goals: it encourages parent participation and tries to improve family functioning and parenting through the use of social workers, home visits and parental education. In addition, it serves as free childcare. This makes parents, and mothers in particular, another group who may be directly impacted by the program.

Due to the breadth of services that Head Start provides, it is theoretically unclear if and how mothers would have changed their behavior under its influence. If it acted primarily as additional years of primary school, it should mirror the labor supply impacts of kindergarten and encourage entry into the labor market, as in Cascio (2009). If the training, volunteering and informational aspects of the program were most impactful it could change mothers' beliefs on parenting, education, employment or welfare and cause changes in their behavior. For example, if Head Start's social workers help families receive more public assistance to which they are eligible, they may experience work disincentives and remain on public assistance and out of work for longer. If the program teaches mothers that being at home with their child is important for development, it may change the cost/benefit calculation of employment and result in dropping out of the workforce.

In this paper I consider Head Start's impacts on the short and long-run labor supply, welfare use and education of mothers. To identify these effects I use a discontinuity in Head Start availability for the 300 poorest counties. To ensure that the neediest children would

[1] State and local spending on Head Start totaled about 180 million.

be served by the program, its creators sent civil servants to the 300 poorest counties to encourage and assist the first round of program application grants (Zigler and Valentine, 1979; Ludwig and Miller, 2007). To the best of my knowledge, this is the first paper to estimate the causal effect of Head Start on long-run maternal outcomes, and one of the first to look at maternal outcomes more generally. This time period is especially compelling because the policy landscape and counterfactual was much simpler than today. Poor mothers were not expected to work to receive public benefits and there were far fewer options for formal child care available to the poor (Cahan, 1989). This allows me to estimate the effects of Head Start without the added complications of interactions with other policies shifting maternal work and childcare options.

Using restricted 1970 Census data I find no impact of the discontinuity in Head Start on employment or welfare outcomes for mothers of preschool aged children as a whole. However, for single mothers, there is a clear *decrease* in labor force participation and hours worked and suggestive evidence of an increase in welfare receipt. In the long run, 10 years after their child was Head Start eligible, these results seem to be compounded. Non-white and un-married mothers work less both five and ten years later. There is also evidence of a return to school for these mothers, and suggestive increases in welfare use for the single mothers. I argue that these results are inconsistent with Head Start acting primarily as free child care, instead they suggest that the program affects mothers most strongly through the transfer of information: their eligibility for public assistance and the importance of close parenting for child development. That the negative impacts on work persist for 10 years implies that Head Start had very powerful impacts on mothers.

Since survey data is often a poor measure of public assistance recipiency (Bound et al., 2001), I also look for evidence of changing welfare usage due to Head Start in an administrative survey of welfare case-files. Since this population is composed solely of welfare-recipients I cannot directly observe entry or exit into welfare. Instead I test for changes in composition: whether the discontinuity in Head Start affected who receives public assistance. I find no

evidence that Head Start pushed new families onto welfare, instead I find that it may have discouraged households from leaving welfare, particularly single-mother families.

To support my reduced form results, I estimate the size of the discontinuity in Head Start availability as measured by dollars of funding and enrollment among children. The funding data shows noisy increases of roughly 100% in dollars per age-eligible child at the cutoff, which peak in 1970. Using Census and AFDC data I find that child enrollment increases by 6-15 percentage points, depending on subgroup. I also calculate that the discontinuity in funding should allow for a larger increase in enrollment than what we observe. This is evidence that the counties that received grant-writing spent more money per child and may have provided more or higher quality services.

As an extension, in the Appendix, I compare these impacts of Head Start to the impacts of Kindergarten. I use a regression discontinuity on birth month to identify the exogenous increase in Kindergarten eligibility for children born in December as opposed to January (modeled on the strategy used in Gelbach (2002) and Fitzpatrick (2012)). Although this cutoff predicts kindergarten attendance very well, there is no change in household welfare participation. This is true both among welfare recipients a whole, as well as for the subset of counties that identify my results for Head Start. This confirms that although Head Start offers something akin to childcare, its impacts on single mothers are very different.

These results show that the Head Start program acted very differently empirically on mothers than traditional school programs and formal child care programs do. This provides context for research into Head Start more generally and its relationship to the social safety net. The effects on maternal behavior and household incomes are informative of a change in home environment for Head Start children, which may be a mechanism through which Head Start impacted their long run outcomes (Currie and Thomas, 1995; Garces et al., 2000; Ludwig and Miller, 2007; Deming, 2009). The findings in this paper suggest that in addition to the services that Head Start provided directly to children, the program increased the amount of welfare income the most disadvantaged children received and decreased the

employment of their mothers during their childhoods. This likely implies that mothers were more likely to be at home with their children between the ages of 4 and 14. Additionally I bring new data to the study of Head Start and update an influential paper in the literature, Ludwig and Miller (2007), bolstering the identification strategy by correcting a calculation error in the categorization of treatment and control counties.

The paper proceeds as follows: In Section 2 I describe the institutional background, the findings of existing research and the intuition behind the identification strategy. In Section 3 I describe the data sources used in estimation. In Section 4 I detail the empirical estimation strategy. Sections 5.1-5.2 present evidence of the existence and size of the discontinuity. Sections 5.3-5.4 present the contemporaneous impacts on mothers, while Section 5.5 considers the long-run effects. In Section 6 I walk through robustness checks of my results. Finally, Section 7 concludes.

2 Background

2.1 Head Start

Head Start was created in 1965 as an extension of President Lyndon Johnson's War on Poverty and was intended to give disadvantaged children a "head start" in their education and reduce the achievement gap. To be awarded funds to open a Head Start program, local agencies needed to write grant applications to the Office of Economic Opportunity (OEO) in the spring of 1965. These local agencies were a mix of Community Action Agencies (the local arms of the War on Poverty), school districts, private non-profits and churches. The 1964 Economic Opportunity Act, which created both the OEO and the Community Action Program/Agencies, explicitly wanted to direct poverty-alleviation programs to individual communities. The goal was to allow the poor themselves to determine what would help them the most. Unlike many federally funded initiatives before or since, the OEO retained full discretion to send money where it chose and without passing through state or local

governments (Bailey and Duquette, 2014).[2]

The first Head Start programs were summer only, running from June-September of 1965. However, from the beginning, the OEO planned to also offer year-round programs. The first of these began in the fall of 1965 and within two years full-year enrollment grew to roughly 200,000. Since none of my outcome data was collected during the summer months, and full-year programs are a more intense treatment, I focus on them.

The Head Start Program was motivated by the desire to prepare poor children for school. However, it didn't do this entirely through education. Programs were intended to provide health, nutritional and psychological services to children, as well as social services, volunteering and educational opportunities to parents. The aspects of Head Start that are intuitively likely to affect mothers were the custodial nature of the program, the social services and the emphasis on parent participation. In the following sections I will detail each of these three components and how they might influence mothers.

Head Start as Child Care

In 1968 full-year Head Start programs were typically nine months long, with an average of five hours of class-time per day. Roughly a quarter of programs were full-day, offering more than six hours per day of supervision (Levitan, 1969). Programs were required to serve children for at least 15 hours per week and eight months per year for full-year programs (Project Head Start, 1967). This means that Head Start provided at least as many hours of care as a typical half-day kindergarten program, and often many more hours. Thus for a woman on the margin of entering the labor force, Head Start provided a wage subsidy equal to the cost of alternate childcare for those hours. This effectively reduced the reservation wage and should have, all else equal, induced an increase in labor force participation on the extensive margin. For mothers who were already working and purchasing child care it also

[2]The OEO, and later the departments of Health Education and Welfare and Health and Human Services, were supposed to allocate funds proportionally to a state's poor population, but within the state they had control over where the money went.

had income effects. By reducing their expenditure on child care, Head Start would increase their available income and this could result in fewer hours worked.

The limited hours of the modal Head Start program may make it difficult to hold full-time employment. The informal care that was more typical for working mothers during the period would have been more flexible than Head Start's classroom structure.[3] If there are frictions in choice of hours to work, Head Start's schedule may complicate employment (Dickens and Lundberg, 1993; Ilmakunnas and Pudney, 1990).

Social Services and Home Visits

According to a 1967 Head Start manual, "the prime responsibility of the social service staff is to plan for and ensure that the children and families get all the services to which they are entitled. Their role should be that of a strong advocate in obtaining services from local agencies and in referring families to them" (Project Head Start, 1967). Teachers as well as social workers made home visits, with the latter targeting three such visits per year (Project Head Start, 1967). These would give representatives of Head Start a one-on-one opportunity to discuss how a mother could be helping their child.

Means-tested welfare programs including Aid to Families with Dependent Children (AFDC) have been shown to depress working behavior (Moffitt, 1992; Hoynes, 1996). Thus, if Head Start made women aware that they were eligible for welfare we would expect a decrease in employment and increase in welfare receipt. While these outcomes would today be generally considered a "failure" of the program it is important to consider the context. Working, especially for mothers of young children, was not encouraged during this period. The leading child development experts of the 1940s and 50s believed that non-maternal supervision, and group care in particular, was psychologically damaging to children (Cahan, 1989).[4] The

[3] Heckman (1974) discusses the effectiveness of subsidized child care programs to affect labor supply when cheap informal care is common.

[4] This view was espoused in the best selling book of the 20th century: The Common Sense Book of Baby and Child Care, by Dr. Benjamin Spock. Much of the motivation for the damaging effects of non-maternal care came from institutionalized orphans and the work of Rend Spitz.

message that these Head Start mothers were likely hearing from their child's teachers was that being an available and attentive (stay-at-home) parent would be good for their child's development.

Volunteering and Participation

Many parents were hired by the centers or encouraged to volunteer there.[5] These parent employees and volunteers where overwhelmingly mothers. Mothers could also be directly hired by centers: 83% of centers in 1968 reported that they had parents among their staff, accounting for 29% of paid staff positions (Project Head Start, 1970). The numbers of volunteers vastly outnumbered paid workers, but information on usual hours worked is limited. Sources seem to indicate that most volunteers worked only a few hours per year at the center, although a minority likely worked much more (Project Head Start, 1970).

Actual employment can be roughly observed in the data, and might either increase or decrease observed hours of work, depending on the counterfactual activity for the women who Head Start would have hired. However, volunteering cannot be measured by the data I use, and if it crowds out for-pay employment, may result in a decrease in measured work.

2.2 Identification Strategy

In early 1965, during the inception of the first Head Start summer program, the OEO was concerned that the poorest parts of the country would be ill-equipped to apply for Head Start grants. These would need to be written in just a few short months to receive a program in summer 1965. To try to ensure that the neediest communities would still benefit from Head Start, the OEO decided to target additional assistance to the 300 poorest counties. This assistance took the form of sending civil servants to each county to identify individuals who could lead a program and help them write a grant (Zigler and Styfco, 2010).

[5]The first summer program hired approximately 100,000 individuals, the majority of which were parents. An additional 100,000 parents volunteered. In 1968 there were about 20,000 paid employees staffing the year round programs, with almost 40,000 volunteers (Levitan, 1969).

This decision is what makes a regression discontinuity approach feasible. Although the assistance in grant writing only occurred in 1965, it is reasonable to expect that the discontinuity's impact on program availability would persist for several years after. First, the agencies that were awarded grants in that first summer were very likely to maintain funding in the subsequent years. In 1970, when the Head Start Bureau director was asked how many of the original programs had been closed, he responded, "Well, Ed there *was* one program we *almost* closed down" (that Boston center had been beating children with sticks) (Zigler and Muenchow, 1992). Since centers were not being shut down, any that were created in the beginning likely persisted in years ahead. Second, Head Start funding was essentially flat from the late 1960's until 1990. This makes it unlikely that many new centers were being opened (which would have attenuated the discontinuity). Section 5.1 shows that the discontinuity persisted at least until 1971, and likely through 1980.

In Figure 1 I summarize the locations of the relevant counties. Panel (a) shows 1960 poverty rates for all counties. It is clear that extreme poverty rates are concentrated in the south. Panel (b) identifies the locations of the 300 poorest counties, which are clustered in states like Mississippi, Kentucky, Texas, Alabama and Georgia.[6] My main analyses consider the difference between outcomes at the county level at the discontinuity.

2.3 Related Literature

A very small literature examines the impacts of Head Start on parents.[7] In the research closest to that of this paper, Sabol and Chase-Lansdale (2015) consider the labor supply and educational decisions of both parents. They use the federally funded, randomized control trial of Head Start attendance (the Head Start Impact Study or HSIS) and find no significant

[6]There are also two county-units in Alaska that were among the 300 most poor, but these are dropped because the administrative boundaries in Alaska changed drastically between 1960 and 1970 (Alaska became a state in 1959).

[7]In concurrently done work, (Long, 2016) uses a similar strategy to that of this paper and covers some of the same ground in contemporaneous impacts on mothers. Our short run results broadly agree, but this paper estimates long-run impacts, focuses on different groups of mothers, uses the corrected running variable, conducts more extensive tests of the discontinuity, more extensive tests of the pre-period and uses different methodologies for the estimation of the RD.

change in work, but do see an increase in educational attainment for parents whose children were randomly assigned to Head Start enrollment at age 3. The lack of labor supply impacts during the 2000's is not surprising given the work incentives built into the welfare system by this time. Additionally the counterfactual to Head Start included many other publicly subsidized preschool and child care options to which the parents would have had access (Kline and Walters, 2014). In contrast, the 1960s and 1970s which I study had neither of these features, and ultimately I find that Head Start had quite strong employment effects during this early period. Additionally, I am able to look at long-run outcomes for moms, which is not possible with the HSIS because data collection was discontinued after the children reached 3rd grade.

Another strand of closely related literature looks at the impacts of publicly provided kindergarten on maternal labor supply. Although kindergarten is universal and Head Start accepts only disadvantaged children, both are educational programs for young children which implicitly provide similar amounts of free child care to parents. Cascio (2009) uses the rollout of Kindergarten programs between the 1960s and 1980s to identify the impact on maternal labor supply. The states she considers are primarily in the south, which is also true of the counties that identify the discontinuity in this paper. She finds a large increase in employment for single mothers whose youngest child is kindergarten age.[8]

Gelbach (2002) and Fitzpatrick (2012) follow similar approaches, using school start age cutoffs and when a child was born to identify the impacts on maternal labor supply. Gelbach (2002) focuses on the 1980s and finds increases in work for both married and single mothers and a decrease in welfare receipt for single mothers without younger children. Fitzpatrick (2012) uses the 2000 Census and finds that increases in employment are only observed for single women without younger children. Taken together these studies show that kindergarten can increase employment, especially for single mothers.

[8]It's reasonable to expect that women whose youngest child becomes eligible for free childcare will have a larger extensive margin increase in employment than mothers who still have younger children at home. That distinction would also be of interest in my setting, unfortunately the sample sizes for single mothers are too small to further divide them by age of youngest child.

Two papers looking at Head Start's impacts on parenting behavior in the HSIS find an increase in the time that parents spend reading to their children (Gelber and Isen, 2013; Ansari et al., 2016), and that parent involvement in the centers decreases spanking and increases time invested in children's cognitive stimulation (Gershoff et al., 2016).

Head Start's impacts on children have also been widely studied. Several papers using quasi-experimental variation from the earlier years of the program have found meaningful impacts on child mortality, long run education and labor market outcomes (Currie and Thomas, 1995; Garces et al., 2000; Ludwig and Miller, 2007; Deming, 2009). The long run impacts follow short run gains in test scores for Head Start children, which fade out by late elementary school (Barnett, 1992). Large positive impacts on children have also been found for non-Head Start preschool programs, such as Perry Preschool and the Abecedarian Project, that were also implemented pre-1980 in the US (Campbell et al., 2002; Masse and Barnett, 2002; Schweinhart et al., 2005; Belfield et al., 2006; Anderson, 2008; Heckman et al., 2010; Campbell et al., 2014), and in Europe (Havnes and Mogstad, 2011; Rossin-Slater and Wüst, 2016). The analysis of the HSIS finds a similar short run pattern of increased test scores immediately after Head Start, which disappear by the time the study ends in 3rd grade (Puma et al., 2005, 2010). Several papers have dived deeper into the HSIS and find that Head Start does benefit the most disadvantaged participants (Bitler et al., 2014), and that Head Start's positive impacts are larger for children who would not otherwise attend preschool (Kline and Walters, 2014; Feller et al., 2016).

Since Head Start serves families which are often eligible for welfare, and the program explicitly encourages connecting families to benefits, there is also a connection to the literature which links information to program take-up. Daponte et al. (1999) find that some families that are eligible for food stamps are not aware of their eligibility. They then randomly share information about food stamp eligibility with families and find a significant increase in benefit use. This implies that lack of information is partially responsible for low take-up rates. Aizer and Currie (2004) look at free pre-natal care use by zip-code and ethnic group. They

find that take-up is highly correlated within neighborhood and ethnic group, but that the likely reason is differences in the institutions serving different groups of low-income women (and not information-sharing).

If Head Start does link families to welfare benefits due to information or changing beliefs about the importance of stay at home parenting, the literature on welfare's work disincentives is also relevant. The 1960s and 1970s were a time of explosive growth in the welfare population and the benefit structure provided strong work disincentives. Before 1968 earnings and income reduced benefit dollars essentially one-to-one. Following a 1968 policy change the first $30 in monthly earned income was disregarded as was a third of additional earnings (Ways and Means Committee, 1998). These high implicit tax rates have large theoretical and empirical work disincentives (Moffitt, 1983, 1992; Hoynes, 1996).

3 Data

3.1 Restricted Census

My primary source of outcome data are the restricted use 1960-1980 long-form Decennial Censuses.[9] My main analyses are based on 37,000 mothers of Head Start aged children in 1970. The restricted use data provide necessary geographic information and a large sample size.[10]

I construct a file of children who are matched to their mothers.[11] There is one observation per child and maternal characteristics of interest are attached to each child observation. These data are used as-is when estimating child enrollment, but since mothers can have

[9]The 1970 Census had two long form questionnaires, but I primarily use the 15% sample. 1960 is a 25% sample of the U.S. population, and 1980 is also a 15% sample.

[10]In public use data, counties are aggregated into county groups with at least 250,000 in population (100,000 in 1980). The poorest counties tended to be very small in terms of population, thus an affected county generally makes up only a small fraction of the county group it's placed in. Further, the restricted samples are at least three times larger than the public samples, which helps with power.

[11]These matches are constructed to mimic IPUMs "momloc". The match includes likely non-biological mothers such as stepmothers. Any children who cannot be matched to a mother in their household are dropped from the sample.

multiple children I drop duplicate mother observations when considering maternal outcomes. My main samples are children born in 1964-1965: these children would have been between the ages of three and five during the beginning of the 1969-70 school year.[12] For maternal outcomes I look at mothers of children with these ages.

For mothers I consider labor supply (employment and hours) the previous week and welfare receipt for the previous calendar year. I also consider school enrollment for mothers. Although the Census is only recorded every 10 years, in addition to asking labor force questions about the past week and year, there were also five-year retrospective questions. Thus I can roughly measure employment levels every five years between 1960-1980.[13] For children, I examine whether their highest attended grade is nursery school.[14] I also look at nursery school and kindergarten attendance combined. I use Census person-weights in the estimation.[15]

3.2 1967 AFDC Survey

I supplement my analyses using data from the 1967 Survey of Aid to Families with Dependent Children (AFDC) case-files conducted by the Department of Health, Education and Welfare: Social and Rehabilitation Service (HHS: ASPE, 2003; DHEW, 2011).[16] This survey was a 5%, stratified, random sample of AFDC households, and covered all states. Information was filled out by the AFDC caseworker to the best of their knowledge about the household (they were instructed to answer "unknown" instead of guessing). Importantly, the dataset includes whether or not each child in the household was attending Head Start

[12] On Census day this includes the oldest 3/4ths of four year olds, all five year olds and the youngest quarter of six year olds. Birth years are constructed from age in years and quarter of birth.

[13] The five year retrospective question is only asked for women who were at least 14 in April of 1965, thus it is missing for a small number of very young moms. To keep the sample consistent across specifications and because the women were so young at the time I assume they were not employed and assign them to not-working.

[14] Census documentation confirms that Head Start should be included in this category. This either means they are currently enrolled in nursery school or they were previously but have not yet begun primary school.

[15] The Census discourages the use of restricted data for tables of means. Appendix Table A2 reports summary statistics using County Data Books (1972 and 1983).

[16] I thank Andrew Goodman-Bacon for sharing these data with me (Goodman-Bacon, 2016).

at the time of survey (November or December 1967).[17] Since survey measures of transfer income have been shown to suffer from large scale misreporting (David, 1962; Goudreau et al., 1984; Bound et al., 2001; Meyer et al., 2015), the administrative nature of this source can be expected to measure the population of welfare receivers more accurately than the Decennial Census.

The surveys are not designed to identify counties, but the documentation includes a "county code" which is in most cases determined by the alphabetical ordering of counties. I discuss the details of matching counties to county codes in Appendix Section B.4, ultimately 95% of individuals in the data are matched to a county. The data include race, birth month and year of all children, location of the father, the work activities of the mother as well as all sources of government assistance and other income sources for the families. I include survey person-weights in the regressions.

These data are summarized in Table 1. Because all individuals in these data are AFDC recipients there is less of an obvious disadvantage gradient compared to the Census data in Appendix Table A2. This table suggests some early evidence for the discontinuity: the children in the counties that are slightly poorer than the cutoff have notably higher Head Start participation rates compared to those below (7.8% compared to 3.4% for 4 year olds).

3.3 Head Start Funding Data

In order to test the size of the discontinuity in Head Start availability I turn to county funding levels. These funding data are available for years 1966-1968 in the Community Action Program (CAP) grant files and for 1968-1980 from the Federal Outlay System files (FOSF).[18] Head Start related grants are isolated from other funding streams and aggregated to the county level. The grants are assigned to the county in which the grantee is headquartered. Head Start programs may operate centers in multiple counties, but their grant would only

[17] Similar surveys were enacted biennially after 1967, but later years are shorter, have less geographic coverage, and are missing a question on Head Start.

[18] I use a version of these data which were cleaned by Esra Kose and Doug Miller. To my knowledge, no grant-level funding data from the first year of the program exits.

appear in one. Since I am ultimately interested in the counties in which children were served, not where programs are headquartered, these data are imperfect. The grant data is a proxy for the amount of funding available per child in a county.[19]

The funding is scaled by the number of three to four year olds in a county-year (Surveillance, Epidemiology, and End Results (SEER) Program, 2013), and thus measures the available funding per Head Start aged child in the county.[20]

3.4 Putnam File

The variable which defined the 300 poorest counties, 1960 county poverty, comes from an even older dataset. This so-called "Putnam File" was a 1964 re-analysis of the 1960 Census created for the OEO. Since the concept of a "poverty-threshold" did not exist at the time of the 1960 Census, this dataset provided OEO with various statistics on poverty by household type at the local level. Although exactly what data was used by the OEO to determine the 300 poorest counties is unknown, this is the mostly likely source.[21] From these data I create "Persons in Poverty" and "Persons", whose ratio is the running variable. I also use this to determine and calculate the cutoff: the poverty rate of the 300th poorest county (59.35%). In this process I uncovered a coding error in Ludwig and Miller (2007) that caused their poverty rate to be slightly off and switched several counties from one side of the cutoff to the other. This process is documented more thoroughly in Appendix Section B.2.

[19]The CAP data includes some information on the counties in which a a grant was spent, but it is not reported for a large number of grants. Allocating funding to the counties that are reported to be served does not yield any improvement in precision.

[20]Since these estimates began in 1969 I extrapolate population backwards to 1966.

[21]The version housed on the NARA website in an outdated computer format that is labor intensive to make usable. I thank Jens Ludwig and Martha Bailey for separately sharing their versions of this file. The two versions were identical, but obtained from NARA in different years.

4 Empirical Strategy

To estimate the reduced form impacts of Head Start availability, I take advantage of the discontinuity created by the decision to target the 300 poorest counties with grant writing assistance. This allows me to identify the difference in outcomes due to this additional assistance. The running variable is 1960 county poverty and the cutoff is a poverty rate of 59.35%. Intuitively the children living in counties that were just poorer than that cutoff became more likely to be served by Head Start. The great benefit of this approach is that it requires only quite weak assumptions to give causal estimates. In this setting the necessary assumption is that in the absence of the grant-writing assistance, the relationship between 1960 county poverty and the outcome variables would be smooth through the cutoff.

The cost of the regression discontinuity (RD) approach is that the impacts are local to the cutoff; i.e. it identifies the added benefits of Head Start to individuals who live in counties near the cutoff. Thus these are estimates which are valid in counties with very high poverty rates, primarily in the south. They are not necessarily externally valid to other parts of the country or other populations.

I implement regressions of the form:

$$Y_c = b_0 + f_1(P_c - P_{300}) + \alpha G_c + G_c f_2(P_c - P_{300}) + \beta X_c + \nu_c \tag{1}$$

$$G_c = 1(P_c > P_{300}) \tag{2}$$

where P_c is the poverty rate in county c, and P_{300} is the cutoff. G_c is an indicator equaling one if county c had a poverty level below the cutoff. $f_1()$ and $f_2()$ are functions of the distance from the cutoff. In practice I estimate these functions using local linear regressions and a triangle kernel for a given bandwidth (Porter, 2003; Cheng et al., 1997). Y_c is the outcome for the relevant group at the county level: in the proxy first stage this is preschool attendance or funding, in the main results it is some maternal outcome. I include a vector

of county level controls, X_c, in my main specifications to increase precision.[22]

This approach yields causal estimates of the impact of Head Start grant-writing assistance as long as potential outcomes are smooth through the cutoff; i.e., nothing else about the counties is discontinuously changing there. This seems intuitively true, there is not likely to be anything inherently different between the counties on either side of the cutoff, and it is testable for observable characteristics (Lee, 2008). Tests, reported in in Sections 6.1-6.2, for discontinuities in demographic characteristics and pre-Head Start outcomes are reassuring that there were no pre-existing differences between the counties, and that none endogenously developed over time (e.g. through migration). Although it is difficult to conclusively prove that no other program used this 300 poorest rule, it is described as being something new by its inventors and despite considerable searching I have found no record of its use for other reasons.[23]

Regression discontinuities will only yield causal estimates when the agents cannot precisely manipulate their location on the running variable. This is not particularly plausible in this setting since the running variable was established several years before the treatment was announced. Nonetheless, I formally test for a discontinuity in the density of counties at the cutoff as suggested by McCrary (2008). The result is shown in Figure 2, and the density is indeed continuous through the cutoff.

I implement the robust confidence intervals and mean-squared-error optimal bandwidth selection process proposed by Calonico et al. (2014b,a) (CCT). However, due to sample disclosure restrictions for restricted Census data, I report a common bandwidth of seven percentage points of county poverty rates throughout my Census results. This bandwidth

[22]The data also allow estimation to be done at the individual level, with standard errors clustered by county. This is equivalent to weighting the county level regressions by sample size in county. These two strategies yield somewhat different interpretations of the $\hat{\alpha}$. The $\hat{\alpha}$'s I report are the impacts of the discontinuity on the county, while the alternative estimation would be the effect on an individual. While my main results are quite similar between these two approaches, I prefer the county level results because the density of counties is more continuous than the density of individuals. I show this in Appendix Section B.1.

[23]I check for discontinuities in other CAP programs in Appendix Table A3. Additionally, in estimates not reported in this paper, I use REIS data from 1969-1975 to confirm there are no discontinuities in county level funding for Social Security, medical spending, Food Stamps, SSI, UI or training programs.

was chosen as a round number near the midpoint of the range of optimal bandwidths, but allows the sample to remain constant for all outcome variables and reduce disclosure risk.[24] The confidence intervals implement bias-corrected inference that is robust to "large" choices of bandwidth, I report the p-values from these confidence intervals in addition to conventional county-clustered standard errors.

In my main estimations, Y_c will be maternal labor supply (participation and hours), welfare receipt and nominal welfare dollars. $\hat{\alpha}$ gives the effect of being a county below the cutoff, assuming that the relationship between maternal labor supply and poverty would otherwise be continuous through the cutoff. For example, if I implement Equation 1 on mothers in 1970 who had a child aged 3-5, $\hat{\alpha}$ would tell me the average difference in the labor supply of mothers of age eligible children at the discontinuity.

Since the treatment of interest is child's Head Start attendance, this is not a sharp discontinuity. This means that the cutoff only increases the probability of being treated, there are treated individuals in the counties above the cutoff and untreated ones below. What I'm interested in is not the impact of simply living being below the cutoff, but the effect of Head Start. In Sections 5.1 and B.3 I show that the discontinuity seems to affect both spending per child and fraction of children served. This means that the change in probability that a child is enrolled does not capture the whole change in Head Start. Thus even if it were possible to perfectly measure Head Start attendance for the children my sample of mothers I may not want to use this as an instrument for maternal outcomes. I can scale the reduced form estimates by the change in funding, which would allow me to express results per dollar per child. However, my measures of this "first stage" are noisily measured (see discussion in Section 5.1) so I use the reduced form as my primary estimate and report results scaled by funding dollars for my main results.

[24] Allowing the bandwidth to vary for every regression creates implicit samples between outcome variables. These implicit samples could be quite small (and therefore risky to disclose) if the bandwidth changes only slightly.

5 Main Results

5.1 The Discontinuity in Funding

Ludwig and Miller (2007) look at two years of funding data: 1968 and 1972 and find moderately significant increases of roughly 90% at the discontinuity. I extend this exercise to consider each year of funding between 1966 and 1980. Additionally, I use a more accurate estimate of number of children as a denominator.[25] Ultimately I find qualitatively similar discontinuities in funding, but since my preferred MSE-optimal bandwidths are considerably smaller than those used by Ludwig and Miller (2007), I have lower precision in many years.[26]

In Tables 2 and 3, I present estimates of the discontinuity in Head Start funding per child. Each estimate is the $\hat{\alpha}$ from a separate regression of the form shown in Equation 1. Standard errors are of the conventional heteroskedasticy robust variety and stars denote conventional significance levels. The "Robust P" gives the p-value following Calonico et al. (2014b). The control mean measures the mean of the outcome variable for the sample who are in counties just too rich to have received assistance.

The top panel of these tables reports the discontinuity in overall Head Start funding dollars per three to four year old child in a county. In Column 1 of Table 2 we see that on average between 1966-1980 the funding dollars increase by $450 real 2016 dollars per child (83%) at the cutoff. The coefficients for each year between 1966 and 1980 are positive and large in magnitude. However, this increase is only significant at the 10% level for the years 1970 and 1971, when the coefficients show increases of $795 and $698 respectively.[27] These increases roughly double the dollars available per child at the cutoff. The significance in 1970 is especially important because that is a Census year, and we can be confident that the discontinuity in funding is meaningful when I measure reduced form outcomes for mothers

[25] Ludwig and Miller (2007) use a scaling of total population in a county as the denominator, I use the SEER estimate of 3-4 year olds in a county.

[26] They use bandwidths between 9-36 percentage points of the poverty running variable, while my data driven bandwidths fall between 3.7-9.7.

[27] The CCT-robust p-values for these years are somewhat larger, but they drop when slightly larger bandwidths are used.

at that time.[28] A possible explanation for the additional precision in 1970 and 1971 is that the denominator, which is an estimate of the number of eligible children, would be most accurate in the years following a Decennial Census.[29]

These funding data are not ideally suited for measuring treatment at the county level. As mentioned in Section 3, this is because they record the county where a program is headquartered, not where it is actually spent. Since a single program often operates Head Start centers in multiple counties, the county level funding data becomes a noisy proxy for actual Head Start availability in a county. While it measures something that should be strongly correlated with what we're interested in, it adds considerable noise. It is also a reason to believe that the true increase in Head Start funding availability could be larger than what we can estimate with the available data.

Panel B looks at the change at the discontinuity for whether a county received any funding. These are insignificant and small for all the years I look at. Panel C looks at the other component of overall funding, dollars per child among the subset of counties that received any funding that year. Here we see a starkly different pattern: on average between 1966-1980 funded counties received an additional $1,606 per child at the cutoff (roughly a 90% increase). In 1970 and 1971 it is well over $2,000. The corresponding figures for the three panels for the entire period and 1970 alone are shown in Figure 3.

There are two ways that the grant writing assistance could have helped counties: the first is by making it more likely that they applied for a program, the second is that they wrote more ambitious grants.[30] The second and third panels of Tables 2 and 3 are consistent with it being the latter. More ambitious grants mean more money, which could be used to serve more children, serve children for more hours per day, or to provide higher quality and more

[28] A subset of my outcomes come from welfare users in 1967, in that year the estimated increase in funding is roughly $400 per child. This is an economically meaningful increase, but statistically insignificant.

[29] SEER population estimates are created using Census data.

[30] We can also ask if the Head Start grant writing assistance spilled over into other CAP programs. In Appendix Table A3 I use data from Bailey and Goodman-Bacon (2015) on non Head-Start CAP funding (administrative, legal services, elderly services and health services) to test this. Overall there is no evidence that the discontinuity affected other CAP programs.

services per child.

5.2 The Discontinuity in Enrollment

Ludwig and Miller (2007) use the 1988 National Education Longitudinal Survey (NELS) to estimate the discontinuity in enrollment. The survey captures 8th graders in 1988 and asks a retrospective question on Head Start attendance (in what would have been the late 1970s). The resulting increase in enrollment at the discontinuity is more than 20 percentage points, nearly doubling the rate of enrollment. However, by comparing the national enrollment implied by the NELS to administrative counts of Head Start enrollees I find over-reporting of 40-100%. I discuss this in further detail in Appendix Section B.3, which also identifies reporting errors in the Census and AFDC data. Large scale over-reporting of Head Start attendance in the NELS suggests that the size of the discontinuity is also over-estimated.[31] In the remainder of this section I investigate the increase in enrollment at the discontinuity using the Census and the AFDC data.

Both the AFDC data and the Census have measures of school enrollment that should capture Head Start. Columns 1-3 of Table 4 use data from the 1970 Census. The first two panels of Table 4 consider the proportion of children born between 1964-65 who have reported ever being enrolled in nursery school, but have not begun any other grade. The mean of this variable in the sample population is roughly 5% (reported as "BW Mean").[32] I report the estimated RD coefficient both with and without covariates, and keep the bandwidth fixed at 7. Among the full population of children enrollment rises a significant 2-2.7 percentage points (roughly 50%). Only a subset of these children would actually have low enough family incomes to be eligible for Head Start. These coefficients are larger for the non-whites and children of single mothers (in Columns 2-3), who had much higher poverty rates.

[31] Patterns of misreporting may or may not be continuous through the RD cutoff, and could also be moderated at the high end by the 100% bound. The NELS shows enrollment rates of roughly 20% below the cutoff and 40% above, so if reported attendance is 100% too high throughout the distribution the true discontinuity would be 10 percentage points instead of 20.

[32] I report the mean for the entire sample to limit the creation of implicit samples in the restricted data.

In 1970 Head Start was still a very young program, and since many centers were run out of existing charities, churches and school districts it's not clear that parents would have understood the difference between Head Start and kindergarten. To attempt to account for this type of misreporting, I widen the outcome measure to include both nursery school and kindergarten for these cohorts, to see if that captures the discontinuity better. Results for this broader measure are reported in Panel C and D of Table 4. These estimates show an increase in enrollment among all children by up to 6 percentage points, among non-white children of 15 percentage points and among children of single mothers of 8 percentage points. Significance fades for the overall group (especially the CCT robust p-value), but the two disadvantaged groups now have much lower p-values.

Figure 4a provides a graphical illustration of this discontinuity for single mothers. It plots the local-linear relationship between enrollment and the running variable, separately on either side of the cutoff.[33] On the left we can see that as counties get poorer, they have lower enrollment, but at the cutoff enrollment discontinuously increases and stabilizes at about 20%. Since there is no reason for true kindergarten availability to differ at the cutoff, I conclude that the combined measure is capturing an increase in Head Start attendance, through misreporting of Head Start as kindergarten.

Column 4 of Table 4 shows the estimated increases in Head Start enrollment for AFDC children born 1962-1963 (representing the same school-age cohorts as previous). The 5-6 percentage point increases (up to 190%) estimates are not significant, even using a larger bandwidth of 12. Results with optimal bandwidth and broken out by year of age are shown in Appendix Table A4.[34] As with the 1970 Census, there is evidence of underreporting of Head Start attendance in the AFDC data. If I broaden the outcome variable to include other

[33]All the sub-figures in Figure 4 use a bandwidth of 7 and a triangle kernel, to mirror the estimated coefficients. The exact size of the discontinuity may differ from the tabular estimates because the latter include control variables. While RD figures customarily include raw means of the outcome against the running variable, the Census prefers to release smoothed representations of data. I provide 95% confidence intervals to give a sense of the underlying variability.

[34]The control means of each birth cohort between 1961-1964 confirm that Head Start attendance is concentrated among just those born in 1962 and 1963.

types of school, the increase is 9 percentage points, for an increase of 155% at the cutoff. This increase is significant at the 5% level using conventional standard errors, but not with CCT-robust inference and not with the inclusion of covariates. The imprecision is relative to the Census estimates is likely due to the much smaller sample size.

Overall, the evidence for a change in Head Start availability at the discontinuity is a 80-100% increase in funding dollars per child, and a 5-6 percentage point increase in attendance for children (not restricting to those eligible). The increases in attendance are higher among AFDC recipients (up to 9 percentage points), non-white (15 percentage points) and children of single parents (8 percentage points). There may also be an increase in the quality of the Head Start program, e.g. the amount spent on services per attendee. In 1968 Head Start spent $1,000 nominal dollars ($6,900 in 2016 dollars) per child in a full year program (Levitan, 1969). This information allows me to do a back-of the envelope comparison of the discontinuity in funding to enrollment. The 1970 and 1971 discontinuities amount to roughly one tenth of the cost of Head Start per child. This implies that the additional funding per child would provide 10 percentage points more enrollment. [35] Since the estimated increase was only 5-6 percentage points, this suggests that the grant-writing assistance may have caused these counties to create programs that were more expensive and potentially provided more services per child.

While the 20 percentage point jump in enrollment reported in Ludwig and Miller (2007) appears to overstate the discontinuity, these many measures of the change in Head Start access point strongly to a real and large first stage. My empirical findings are compounded by the historical evidence on these events and supported by the robust findings of improvements in child mortality reported in Ludwig and Miller (2007) and confirmed by Cattaneo et al. (2016).

[35] Since the number of children denominator includes all children, the enrollment effect should be larger among poor children.

5.3 Contemporaneous Impacts on Mothers

I now focus attention on the effect of the discontinuity on mothers of children who are Head Start age, using the 1970 Census. Consistently with previous sections, I consider mothers who have children born 1964-1965. I look first at employment in April 1965, this is a good placebo test because it predates the first Head Start program by several months. This is shown graphically in Figure 4b. Then I turn to employment and hours worked in the previous week to capture changes in labor force behavior. I also consider whether households received any welfare income as well as the dollar value of welfare benefits in the previous year. I finally examine whether mothers are more likely to be enrolled in school at the discontinuity.

Panel A of Table 5 shows single mothers, a group which is very likely to be eligible for Head Start and the literature has shown is sensitive to the price of childcare. I report only results with covariates included for brevity, but the results are very similar without their inclusion and are reported in Table A5. For these mothers I find *decreases* of seven percentage points in their employment as well as a 3.2 hour decline in hours worked per week. If scaled by the increase in funding I estimated in Section 5.1, this corresponds to a decline in employment among single mothers of 8.7 percentage points for each additional $1,000 dollars of funding per child in the county. However, Figure 4c shows that the change is largely driven by slope changes near the cutoff. While if compared to employment among the same women in Figure 4b, it still seems that there is a decline, but the true effect is likely smaller. The estimated increases in welfare receipt are positive and large in magnitude, but not significant. The results for non-white mothers are shown in Panel B. This group saw the largest increase in enrollment for their children, but there are no statistically significant changes in work, welfare or schooling. A possible explanation for this is that married non-white mothers were less likely to be eligible for public benefits to offset lost income from work.

Panel C considers the entire population of mothers of age-eligible children. At this

aggregate level there are no significant changes in employment or welfare. For mothers as a whole there is no evidence of changed working or public assistance participation. Without the addition of covariates there is a near-significant increase in the average amount of welfare received, and there is suggestive evidence of a small increase in school enrollment for mothers.

Overall these findings show that Head Start had no impact on most mothers while their child was age-eligible. However, for the subgroup of unwed mothers, there were large decreases in labor supply. There is also a suggestive increase in welfare income among this group, which I revisit in the following section. For single mothers these results suggest that changes in information about welfare have played a role in reducing labor supply. However, decreases in work are also consistent with an increase in the perceived value of being at home with children or in-center volunteering crowding out employment.

5.4 Changes in Welfare Use: Administrative Survey

Survey data is known to suffer from misreporting of public program use (David, 1962; Goudreau et al., 1984; Bound et al., 2001; Meyer et al., 2015). This is likely to bias the RD estimates in the Census data towards zero. The 1967 survey of AFDC case-files comes from administrative listings of welfare recipients and therefore should not suffer from misreporting of welfare receipt. Thus I turn to these data to further check whether there were changes in welfare participation by mothers of Head Start eligible children.

Since the data includes only those who are currently using AFDC, I cannot observe direct entry or exit into participation. Instead, I look at the composition of families to estimate if there are new participants or changes in exits. In particular I consider various measures of how long the family has been on the program (length in months, indicators for less than four months, less than 12 months and below median) as well as race, years of education of the mother and whether the father is present in the household. If many families are encouraged to join AFDC as a result of the discontinuity in Head Start, then the proportion of households who began their spell on welfare recently will be discontinuously higher. Likewise, if families

that would otherwise leave AFDC are encouraged to remain on the program the proportion of longer spells will increase at the discontinuity.

The results are reported in Table 6. This table looks at households with a child born in 1961-1962, the same Head Start age group considered previously. The combined evidence on spell length (Columns 1-4) is consistent with longer spells. The share of spells begun recently decreases and the share of above median spells increases. Because I can only look at the composition of spell lengths, the opposite signs of Columns 2 and 3 as compared to 4 are expected. If there are more long spells in the population, the share of those with low spell lengths must decline. Broadly speaking, this increase is consistent with a number of changes in the distribution of welfare recipients, including a decrease in new entrants, an increase in exits among recent entrants and a decrease in exits among those with longer spells.[36] While I cannot fully separate these possibilities, the overall pattern of evidence, including the Census results, is more consistent with the latter explanation. Head Start could have helped mothers who were already receiving some benefits to apply for more forms of support.

In Columns 5-7 I look at the demographic characteristics of families participating in AFDC. If there is a change in who uses the program, we may see it in these demographic variables. Column 5 shows that there appears to be no change in the racial composition of families, and Column 6 shows the same for the years of education of the mother. However, Column 7 provides evidence of a sharp decrease in the probability that the father is in the household.[37] This decrease in the share of families with a father present means that a larger share of recipients are single mothers. These results reinforce the findings of increases in welfare incomes among single mothers in the Census. They are consistent with single-mother households, who would otherwise have left welfare when their child was Head Start age, instead remaining on the program.

[36] If a mother of a four year old would have left welfare when her child was 4 years old, but being connected to more benefits or being informed of the importance of parenting convinced her to stay home - average spell lengths would increase.

[37] AFDC can serve two-parent households when the father is either incapacitated (13% of households) or unemployed (6% of households).

5.5 Long Run Impacts on Mothers

I have established that Head Start had immediate impacts on single mothers, but no observable impact on most other mothers. However, we may also expect that Head Start could have lasting impacts on these women. For the single mothers whose work and welfare use may have changed there may be persistence of these behaviors over time. For mothers more generally, Head Start's informational aspects may influence choices well into the future. To test for long-run impacts on mothers I consider the same cohorts of children as previously, but measure their outcomes in the 1980 Census. These children would have been Head Start age in 1970, but we are looking at their mothers 10 year later (when the children are 14-16).[38] I also report changes in marital status and number of children born, since with 10 years to adjust, these demographics may also respond to Head Start over the long-run.

Panel A of Table 7 presents these long-run estimates for single mothers. Ten years after the results shown in Table 5, the employment rate for these mothers remains lower. This is represented graphically in Figure 4d, which shows that while employment among single mothers is decreasing in county poverty, there is a clear level-shift in the estimated relationship at the cutoff. The coefficient on employment five years ago is also similarly negative, but not statistically significant. Unlike fixed demographic groups like race, the composition of single mothers may have changed (through marriage and divorce). In fact, it did change: in 1970 single mothers made up 13% of the sample, and in 1980 they make up 20%. That means that this population is different from the one measured in 1970. Still, these results could be appropriate for the population of mothers who are single by 1980 unless the discontinuity is also influencing marriage rates.

We can examine this assumption in Column 1 of Panels B and C. In Panel C we find no change in marriage rates for women overall, but for non-white mothers in Panel B, we find an eight percentage point increase in the probability of being married. This tells us that

[38] I focus on this cohort because the discontinuity in Head Start funding is sharpest in 1970 and this is the only cohort of mothers for whom I can observe both proxy measures of Head Start attendance and long run outcomes.

there is also a change in the composition of Panel A due to Head Start, namely that there are fewer non-white single mothers in the counties treated by grant-writing assistance.

If we consider the results in Panel A at face value, we see that employment has been lower both in 1975 and in 1980 (again while statistically significant using conventional methods, these estimates are not significant using the CCT robust p-value) and hours are lower. There also seems to be an increase in extensive margin welfare receipt and a 15% increase in dollars of welfare received. There is also evidence that these mothers are more likely to be enrolled in school, although not that their aggregate years of education have grown over the 10-year period. These findings are still interpretable as the overall impact of Head Start on the population of women who are single mothers by 1980, however given changes in marriage among non-whites it's likely a combination of composition changes and reductions in labor supply.

Panel B shows the coefficients for non-white mothers. Although there were no significant effects on this group while their children were eligible for Head Start, they now look quite similar to single mothers. There are decreases in both employment last week and hours worked last week. The impact on employment is also shown in Figure 4e, which suggests that in counties that weren't treated by the grant-writing assistance, employment rates were steady above 50%, but among the three hundred poorest they were closer to 45%. These large declines are somewhat puzzling since we saw no contemporaneous changes in employment. One potential explanation is that the increase in marriage rates among this population is allowing mothers to stay home. If we translate this into the impact of additional funding on non-white mothers we find that a $1,000 increase in funding per child increases long-run marriage rates by 10.5 percentage points and decreases employment rates by 8.8 percentage points. As with the single mothers there is an increase in school attendance and no change in educational attainment to date. Finally, all mothers are shown in Panel C. Unsurprisingly, since there weren't short run impacts on this group, there seem to be no effects 10 years later.

For the estimates on extensive margin employment, the coefficients are significant at the 10% level using conventional methods, but not significant using CCT robust p-values. However, the decreases in hours worked are more precisely estimated. This evidence of long-run declines in work is likely not something that was intended by Head Start's designers. Possible explanations include that the mothers learn that they prefer being at home with their children (they have more than three on average), and this lasts at least until all their children leave the household. It certainly appears that welfare income is also increasing, so they may prefer the public assistance they're receiving to the low-wage work they would do otherwise. It's not clear if they are better or worse off than they would be in the absence of Head Start. Also the positive effect on schooling seems to imply that they are investing in their educations and may be preparing to re-enter the labor force in the future.

6 Robustness and Specification Checks

6.1 Placebo Tests of the Pre-Period

My identification strategy assumes that there are no discontinuous differences in populations on either side of the cutoff before the grant writing assistance in 1965. One way to verify this is to look at my main outcomes as measured in 1960, five years before the program. Here I consider placebo tests of my main outcomes in the 1960 Census and show that before Head Start was established, the patterns of child enrollment and maternal work did not exist. This in the spirit of the tests suggested by Lee (2008).

In Table 8 I look at kindergarten attendance among children born between 1954 and the first quarter of 1955.[39] I find that there are no significant differences in kindergarten enrollment at the cutoff before Head Start was created. The point estimates are also the opposite sign of the impacts on enrollment in 1970, so any existing differences between

[39]This change in outcome variable and sample is unavoidable since the 1960 Census did not ask about nursery school and only asked any type of schooling question for children who had turned five by April 1st, 1960.

counties would have biased my results towards zero.

In Table 9 I perform a similar exercise on mothers of children born 1954-1955 in the 1960 Census.[40] I find that there are no significant differences at the cutoff for any of the populations I consider. This table tells us that there were not pre-existing discontinuities in the behavior of mothers of Head Start aged children in the two sets of counties. Again, the point estimates among single mothers, are close to zero but also of the opposite sign to the effects that I find in Table 5.

With the AFDC data, I only have one period. As an alternative I look at households that do not contain a child that could have attended Head Start. I remove all families with a child aged 3-8 (the older children could have attended between 1965 and 1967) and consider the same compositional outcomes as in Table 6. These are reported in Appendix Table A11. While the overall spell length shows a marginally significant negative effect, it moves in the opposite direction compared to the treated households (those with a Head Start aged child). There is no evidence of compositional changes in spell length as measured by Columns 2-4, and importantly there is no difference in the share of households with fathers. This tells us that the increase in mother-headed households on welfare is confined to households whose child could have been served by Head Start.

6.2 Continuity of Fixed Demographics

A related set of placebo tests confirms that outcomes or characteristics that should not be influenced by the discontinuity are indeed continuous through the cutoff. In this spirit I test that my main samples of mothers do not differ in their age in years, their racial composition, their marriage rates, the number of children they've born and the probability that their eligible child was born in the first quarter of the year.[41]

Table A9 reports these results. Single mothers, the group for whom I find labor supply

[40] The 1960 Census does not separate welfare income and does not have a five year retrospective work question.

[41] The first quarter variable also confirms that results would not be different if I used years of age instead of birth cohorts to define the sample.

decreases during this period, do not differ at the cutoff in any demographic characteristic. The same is true for the non-white mothers for whom I find long-run declines in employment in 1980. Among the group of all mothers there are two covariates, age and number of children, that appear significant, however this is only true of conventional standard errors. Holding the results to the higher standard of the robust confidence intervals suggested in Calonico et al. (2014b) the p-values are 0.45 and 0.29 respectively.

Endogenous migration is also a potential threat to identification, and one that can be at least partially addressed through this type of test. If Head Start services were so attractive as to incentivize families to move to counties were it was available, then the discontinuity would not be a valid source of identification. If there had been migration in response to the policy we would expect to see differences in the demographics of the counties, those that move would probably be observably different from those that stayed. Table A9 also shows that this was not the case in 1970, there is no evidence that, for example, single moms with more children moved across county lines in response to the Head Start grant writing assistance. For 1980, I use public-use Census data to confirm that the population didn't grow or shrink discontinuously at the cutoff (Appendix Table A10).

6.3 Placebo Cutoffs

Another useful exercise to validate the RD design by testing different cut points: generally we shouldn't see effects for cutoffs that are not the "correct" one.[42] Appendix Figures A1-A2 show a plot of the RD coefficient and 90% confidence interval as the cutoff changes. In Appendix Figure A1 the outcome is the 1970 county Head Start funding per poor child, in Appendix Figure A2 it is the same funding without any of the counties that receive zero dollars ("intensive" funding).

Since even the main coefficient isn't significant in many of these regressions, these are only suggestive evidence. Additionally, since the density of counties decreases sharply to the

[42] Due to Census Bureau disclosure restrictions I do this only with the Funding and AFDC data.

right of the cutoff, the estimates increase steadily in noise. Reassuringly in the majority of the figures, the coefficient that uses the correct cutoff is the largest; in the remainder it's the second largest.

Figure A3 shows the same procedure for the main findings in the composition of welfare recipients: that households with a Head Start age child have more above median spells and are less likely to have the father in the household. Like the results for funding these are quite noisy, but show that most other cutoffs yield estimates closer to zero.

6.4 Choice of Bandwidth

Given the relatively small sample of children, the bandwidth choice matters to the precision of the estimates. In Appendix Figures A4 and A5 I show the effect of changing the bandwidth on the estimate of the discontinuity in funding dollars per child. These are the estimates that I first discussed in the first and third panels of Appendix Tables 2 and 3. I re-estimate $\hat{\alpha}$ for each integer bandwidth between 1 and 25 and plot the estimate and the 90% confidence interval. For brevity I show the entire period mean, and the years whose data I focus on: 1967, 1970 and 1980. These show the same patterns of significance as in the tables, but also reassure us that the level of the estimate is not sensitive to the choice of bandwidth.

Appendix Figure A6 repeats this process for my estimates of the discontinuity in attendance in the AFDC data.[43] As in the forth column of 4, the sample is children born 1962-1963. Head Start enrollment and school enrollment show much the same pattern, but school enrollment is shifted up enough to be significant for bandwidths ranging from 9-16 percentage points of 1960 poverty. Finally, Appendix Figure A7 shows sensitivity to bandwidth for the findings on the composition of AFDC recipients (as in Table 6). While the estimated impacts on spell-length fade at large bandwidths, the decrease in the presence of a father in the home is quite persistent. These emphasize the increase in the share of welfare

[43] Analogous figures are not provided for the restricted Census data because each new bandwidth constitutes a new sample for disclosure purposes.

recipients who are single mothers due to Head Start.

7 Conclusion

In this paper I have estimated the short and long run impacts of the Head Start program on the mothers of eligible children. My identification relies a discontinuity in grant-writing assistance that was provided to the 300 poorest counties when the Head Start program was implemented in 1965. Since counties had no control over whether they met this criterion, it generated an exogenous increase in the availability of Head Start that persisted for years. I leverage the large samples and geographic precision available in the restricted Decennial Census. Using this variation I find that, on average, single mothers were seven percentage points *less* likely to be employed and worked three hours less per week while their child was Head Start age. I also find suggestive evidence that this decrease in employment was accompanied by a 13% increase in average welfare dollars received. Using a 1967 random sample of AFDC case-files I estimate more nuanced changes in welfare receipt. By considering the composition of AFDC families I find an increase in the share of welfare households without a father, and longer average spell lengths. This is consistent with welfare use among single-mothers with a Head Start age child increasing through a drop in the number of families leaving the welfare program. Thus the discontinuity in Head Start would mean that single-mother households that would otherwise have moved off welfare stayed with the program.

Additionally, I provide the first estimates of the long-run impacts of the program on mothers. In the 1980 Census I look at mothers whose children were Head Start eligible 10 years prior. I find that both non-white mothers and single mothers have lower employment at the cutoff and that especially the latter group seems to be using welfare at higher rates. These findings are important because they inform on the uniqueness of Head Start as an intervention, and how it affected the entire household for at least a decade.

Despite its superficial resemblance to kindergarten or pre-k, I've found that Head Start has a different impact on maternal employment. In Appendix Section A I use the AFDC data and school age entry rules to show that public school eligibility does not yield the same effects as in Section 5.4. Head Start does not affect women in the same way that childcare or regular school has been shown to. Many of the features, services and goals of Head Start differ from those of kindergarten in ways that are intuitively likely to depress labor supply. While both have similar classroom features, Head Start also has services outside the classroom and demands on parents' time. Head Start strongly encourages parents to volunteer in the centers, which would take away from potential working hours. Additionally Head Start in the 1960s and 1970s explicitly attempted to link families to any public benefits for which they were eligible, both through information and through direct advocacy in the welfare system. The motivation to create Head Start was largely because of research that indicated that children were particularly malleable at preschool age, and that careful intervention could yield long-run improvements in their cognitive skills and education. Head Start also worked to communicate this sentiment to parents through home visits and lectures on child development. If this caused mothers to believe that spending time at home and interacting with their children was more valuable it also could have led to reduced employment. The mechanisms that act through information like welfare use and parenting are likely to persist beyond when a child leaves Head Start. The long-run impacts that I find are consistent with such persistence.

These results may also be informative in interpreting Head Start's impacts on children. In addition to the health improvements and short run test score gains found by other studies, I show how Head Start likely affected the home environment of non-white children and children of single mothers for the next decade. The long-run decrease in work for disadvantaged mothers means that they would more likely have been at home with their child, and the gains in welfare income tell us that the households had access to more public assistance. Head Start changed more than just what the child learned in the 1-2 years in which they enrolled,

it also affected their home environment long after. This raises the question of whether these changes in maternal behavior and sources of household income are mechanisms through which the long-run findings on children are realized. Although this paper cannot speak directly on whether this was the case, this is an important question for further research.

References

Aizer, A. and Currie, J. (2004). Networks or neighborhoods? Correlations in the use of publicly-funded maternity care in California. *Journal of Public Economics*, 88(12):2573–2585.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*.

Ansari, A., Purtell, K. M., and Gershoff, E. T. (2016). Parenting Gains in Head Start as a Function of Initial Parenting Skill. *Journal of Marriage and Family*.

Bailey, M. J. and Duquette, N. J. (2014). How Johnson Fought the War on Poverty: The Economics and Politics of Funding at the Office of Economic Opportunity. *The Journal of Economic History*, 74(02):351–388.

Bailey, M. J. and Goodman-Bacon, A. (2015). The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. *American Economic Review*, 105(3):1067–1104.

Barnett, W. S. (1992). Benefits of Compensatory Preschool Education. *The Journal of Human Resources*, 27(2).

Barnett, W. S., Friedman-Krauss, A. H., Gomez, R. E., Horowitz, M., Weisenfeld, G., Clarke Brown, K., and Squires, J. H. (2016). State of Preschool 2015. Technical report, National Institute for Early Education Research.

Belfield, C. R., Nores, M., Barnett, S., and Schweinhart, L. (2006). The High/Scope Perry Preschool Program: CostBenefit Analysis Using Data from the Age-40 Followup. *Journal of Human Resources*, XLI(1):162–190.

Bitler, M. P., Hoynes, H. W., and Domina, T. (2014). Experimental Evidence on Distributional Effects of Head Start. *NBER Working Paper*, (w20434).

Bound, J., Brown, C., and Mathiowetz, N. (2001). Chapter 59 - Measurement Error in Survey Data. In Leamer, J. J. H. a. E., editor, *Handbook of Econometrics*, volume 5, pages 3705–3843. Elsevier.

Cahan, E. D. (1989). *Past caring: A history of U.S. preschool care and education for the poor, 18201965*, volume xi. National Center for Children in Poverty, New York, NY, US.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). Early Childhood Investments Substantially Boost Adult Health. *Science*, 343:1478–1485.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., and Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science*, 6(1):42–57.

Cascio, E. U. (2009). Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools. *Journal of Human Resources*, 44(1):140–170.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016). Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality. *Journal of Policy Analysis and Management*, page forthcoming.

Cheng, M.-Y., Fan, J., and Marron, J. S. (1997). On automatic boundary corrections. *The Annals of Statistics*, 25(4):1691–1708.

Currie, J. and Thomas, D. (1995). Does Head Start Make a Difference? *The American Economic Review*, pages 341–364.

Daponte, B. O., Sanders, S., and Taylor, L. (1999). Why Do Low-Income Households not Use Food Stamps? Evidence from an Experiment. *The Journal of Human Resources*, 34(3):612–628.

David, M. (1962). The Validity of Income Reported by a Sample of Families Who Received Welfare Assistance during 1959. *Journal of the American Statistical Association*, 57(299):680–685.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3):111–134.

Department of Health and Human Services (2015). Head Start Program Facts. Fiscal Year 2015. Technical report.

Department of Health, Education, and Welfare. Social and Rehabilitation Service. National Center for Social Statistics. Population Surveys Branch. (2011). Surveys of Recipients of Aid to Families with Dependent Children, 1967-1977. Technical report, National Archives and Records Administration. Social Security Agency.

Dickens, W. T. and Lundberg, S. J. (1993). Hours Restrictions and Labor Supply. *International Economic Review*, 34(1):169–192.

Elder, T. E. and Lubotsky, D. H. (2009). Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers. *Journal of Human Resources*, 44(3):641–683.

Feller, A., Grindal, T., Miratrix, L. W., and Page, L. C. (2016). Compared to What? Variation in the Impacts of Early Childhood Education by Alternative Care-Type Settings. SSRN Scholarly Paper ID 2534811, Social Science Research Network, Rochester, NY.

Fitzpatrick, M. D. (2012). Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool. *Journal of Human Resources*, 47(3):583–612.

Garces, E., Thomas, D., and Currie, J. (2000). Longer Term Effects of Head Start. Working Paper 8054, National Bureau of Economic Research.

Gelbach, J. B. (2002). Public Schooling for Young Children and Maternal Labor Supply. *The American Economic Review*.

Gelber, A. and Isen, A. (2013). Children's schooling and parents' behavior: Evidence from the Head Start Impact Study. *Journal of public Economics*, 101:25–38.

Gershoff, E. T., Ansari, A., Purtell, K. M., and Sexton, H. R. (2016). Changes in parents spanking and reading as mechanisms for Head Start impacts on children. *Journal of family psychology*, 30(4):480.

Goodman-Bacon, A. (2016). Appendix 1 of: The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. Working Paper 22899, National Bureau of Economic Research.

Goudreau, K., Oberheu, H., and Vaughan, D. (1984). An Assessment of the Quality of Survey Reports of Income From the Aid to Families With Dependent Children (AFDC) Program. *Journal of Business & Economic Statistics*, 2(2):179–186.

Havnes, T. and Mogstad, M. (2011). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.

Heckman, J. J. (1974). Effects of Child-Care Programs on Women's Work Effort. *Journal of Political Economy*, 82(2):S136–S163.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1):114–128.

Hoynes, H. W. (1996). Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP. *Econometrica*, 64(2):295–332.

Ilmakunnas, S. and Pudney, S. (1990). A model of female labour supply in the presence of hours restrictions. *Journal of Public Economics*, 41(2):183–210.

Kline, P. and Walters, C. (2014). Evaluating public programs with close substitutes: The case of Head start. Technical report, Mimeo.

Laughlin, L. and Davis, J. (2011). Who's in Head Start? Estimating Head Start Enrollment with the ACS, CPS and SIPP. *US Census Bureau, Working Paper*, 15.

Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675–697.

Levitan, S. A. (1969). *The Great Society's Poor law; a new approach to poverty*. Johns Hopkins Press, Baltimore.

Long, C. (2016). Introduction of Head Start and Maternal Labor Supply: Evidence from a Regression Discontinuity Design. Working Paper 16-35, Center for Economic Studies, U.S. Census Bureau.

Ludwig, J. and Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *The Quarterly Journal of Economics*, 122(1):159–208.

Masse, L. N. and Barnett, W. S. (2002). A Benefit Cost Analysis of the Abecedarian Early Childhood Intervention. In Levin, H. M. and McEwan, P. J., editors, *Cost-effectiveness and Educational Policy*. Eye on Education.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Meyer, B. D., Mok, W. K. C., and Sullivan, J. X. (2015). Household Surveys in Crisis. *The Journal of Economic Perspectives*, 29(4):199–226.

Moffitt, R. (1983). An Economic Model of Welfare Stigma. *The American Economic Review*, 73(5):1023–1035.

Moffitt, R. (1992). Incentive Effects of the U.S. Welfare System: A Review. *Journal of Economic Literature*, 30(1):1–61.

Office of Economic Opportunity (1975). 1960 Census Putnam File.

Porter, M. (2003). The economic performance of regions. *Regional studies*, 37(6-7):549–578.

Project Head Start, editor (1967). *Head Start Child Development Program: A Manual of Policies and Instructions*. Bureau of Head Start and Early Childhood, Office of Child Development.

Project Head Start, editor (1970). *Project Head Start 1968: the development of a program*. Bureau of Head Start and Early Childhood, Office of Child Development, Washington.

Puma, M., Bell, S., Cook, R., Heid, C., and Lopez, M. (2005). Head Start Impact Study: First Year Findings. *Administration for Children & Families*.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., and others (2010). Head Start Impact Study. Final Report. *Administration for Children & Families*.

Rossin-Slater, M. and Wüst, M. (2016). Are Different Early Investments Complements or Substitutes? Long-Run and Intergenerational Evidence from Denmark. *Mimio*.

Sabol, T. J. and Chase-Lansdale, P. L. (2015). The Influence of Low-Income Children's Participation in Head Start on Their Parents Education and Employment. *Journal of Policy Analysis and Management*, 34(1):136–161.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., and Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool study through age 40*. High/Scope Press.

Surveillance, Epidemiology, and End Results (SEER) Program (2013). Populations - Total U.S. (1969-2014) - County-Level Population Files - Single-year Age Groups. Technical report, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch.

U.S. Department of Health and Human Services: Assistant Secretary for Planning and Evaluation (2003). AFDC Data Archive.

Ways and Means Committee (1998). Section 7. Aid to Families with Dependent Children and Temporary Assistance for Needy Families (Title IV-A). *Green Book, United States Government Printing Office Online via GPO Access*.

Ways and Means Committee (2008). Green Book: Head Start. Technical report, House Ways and Means Committee.

Zigler, E. and Muenchow, S. (1992). *Head Start: The Inside Story of America's Most Successful Educational Experiment*. Basic Books.

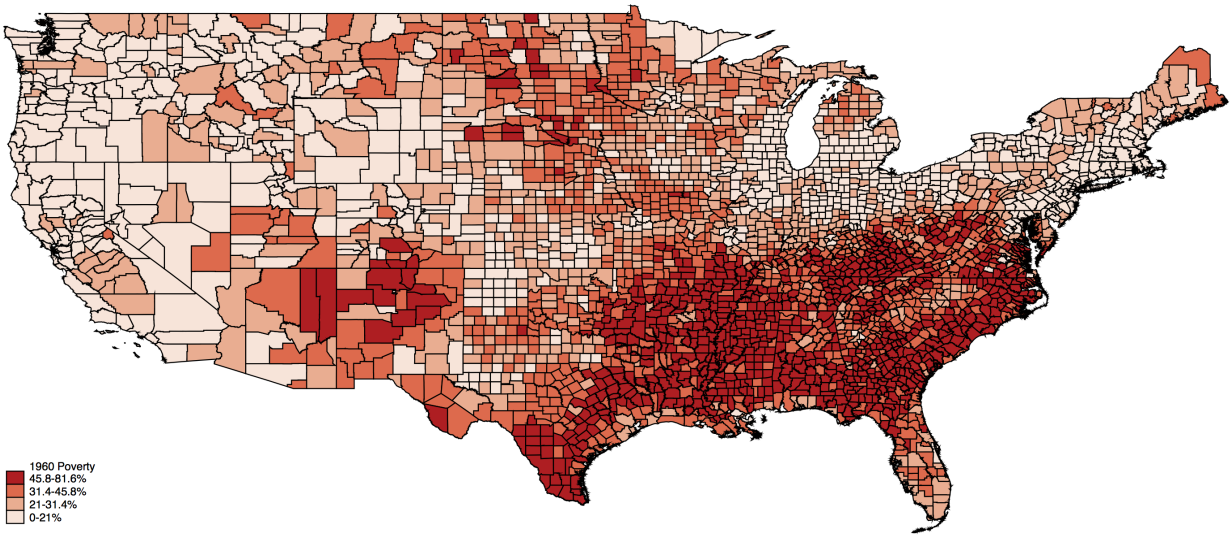
Zigler, E. and Styfco, S. J. (2010). *The Hidden History of Head Start*. Oxford University Press.

Zigler, E. and Valentine, J. (1979). *Project Head Start: a legacy of the War on Poverty*. Free Press.

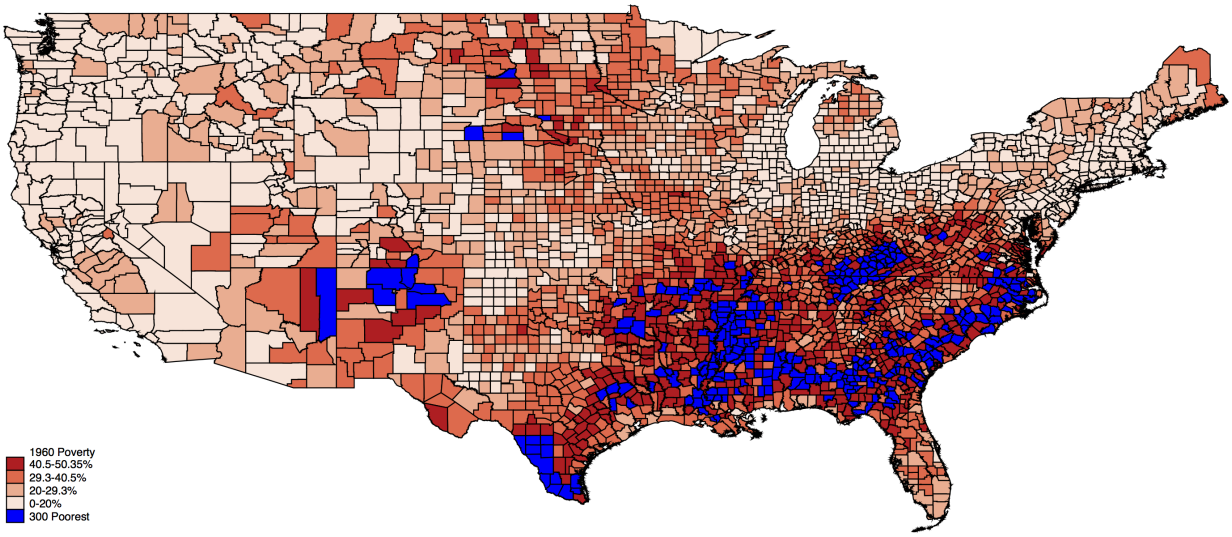
8 Figures

Figure 1: 1960 County Poverty and the 300 Poorest Counties

(a) 1960 County Poverty Rates

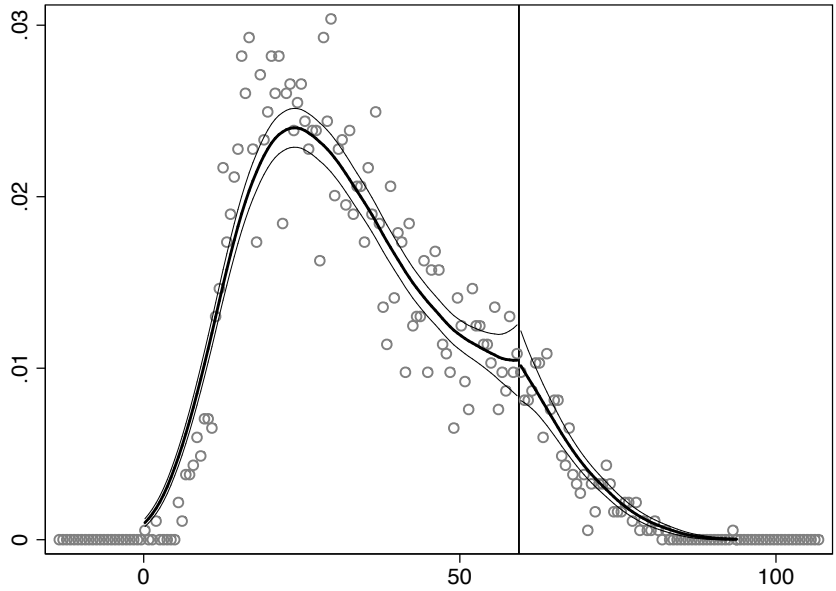


(b) 300 Poorest Counties in 1960



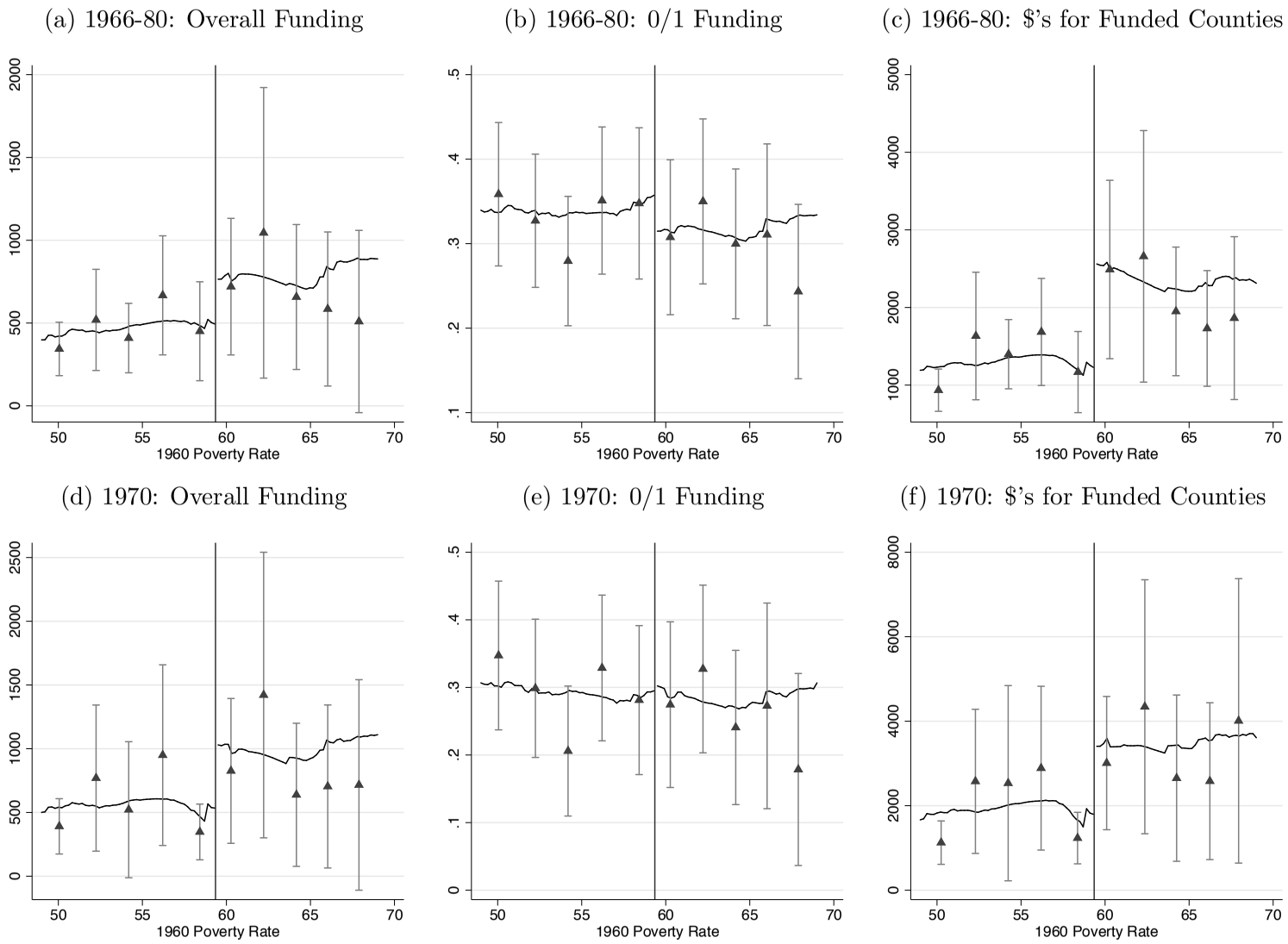
Note: Alaska contains 2 counties within the 300 poorest but is omitted from these figures for simplicity and future analyses because of changing borders of their administrative regions.

Figure 2: McCrary Test of Density of the Running Variable



Note: Figures created using Stata add on "DCdensity" written by Justin McCrary. Data is county level poverty rates in 1960 from the Census Putnam File.

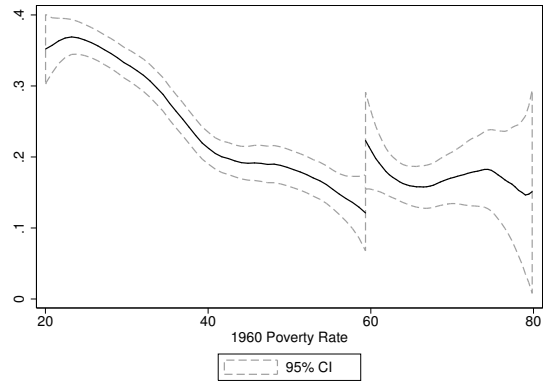
Figure 3: "First Stage" on Funding Per Child



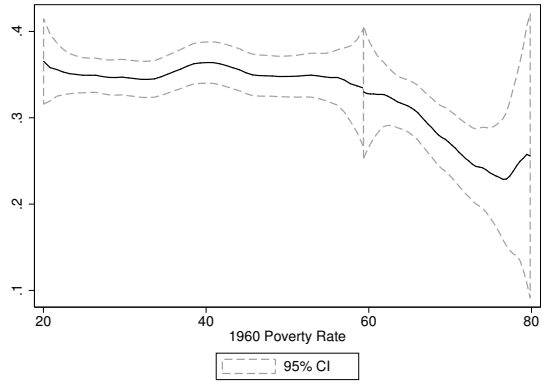
Note: A local linear polynomial is plotted over the interval. The county-level data are shown in 5 bins, along with confidence intervals. The bandwidth used in each graph comes from an optimal bandwidth calculation and matches that used in Table 2. Alaska is not included. 1966-68 funding comes from the Community Action Program files, 1970 funding comes from the Federal Outlay System Files. Funding dollars are scaled by number of 3-4 year olds in a county as estimated by SEER. The last column contains only counties who received positive funding. For the combined year figures the data is first collapsed to county level so that the standard error shown are clustered at the county level.

Figure 4: Census Data Regression Discontinuity Figures

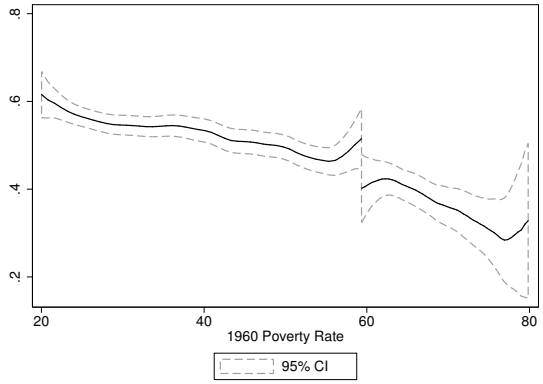
(a) Children of Single Mothers: Enrolled in Nursery School or Kindergarten



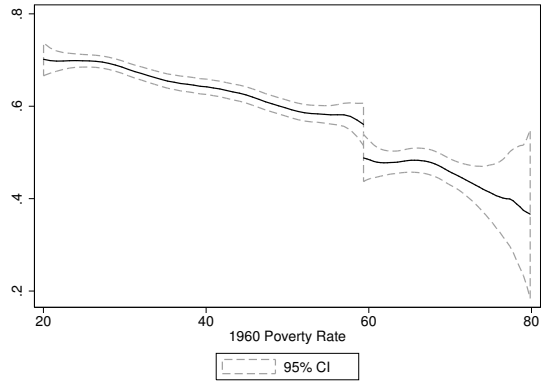
(b) Single Mothers: Worked in 1965



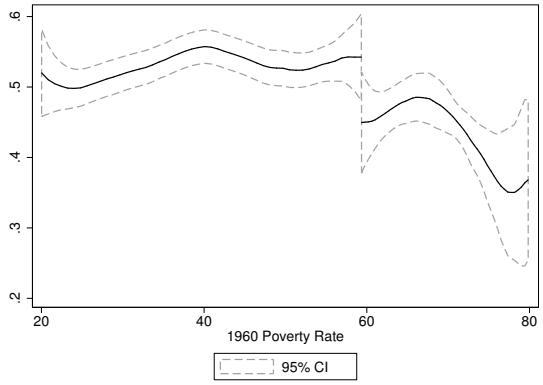
(c) Single Mothers: Working in 1970



(d) Single Mothers: Working in 1980



(e) Non-White Mothers: Working in 1980



Note: Figures plot local linear polynomials of outcome variable (at county level) against 1960 county poverty rate, along with a 95% confidence interval. Triangle kernel, bandwidth is 7 pp of 1960 county poverty. Panel (a) uses all children, born 1964-1965 with single mothers in 1970. Panels (b) and (c) use their mothers. Panels (d) and (e) use single and non-white (respectively) mothers of children born 1964-1980 in the 1980 Census.

9 Tables

Table 1: 1967 AFDC Survey Summary Statistics by County's 1960 Poverty Rate

	All Counties		1960 Poverty 49.35-59.35%		1960 Poverty 59.35-69.35%	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
Child Level Variables						
Age of Child	8.5	(5.1)	9.8	(5.1)	9.7	(5.1)
In Head Start (age 4)	0.1	(0.31)	0.034	(0.18)	0.078	(0.27)
In Head Start (age 3-5)	0.064	(0.25)	0.051	(0.22)	0.071	(0.26)
In School (age 5-17)	0.82	(0.38)	0.86	(0.35)	0.86	(0.35)
AFDC Household/Case level variables						
Num. Children under 6	1.1	(1.1)	0.81	(1.1)	0.86	(1.1)
Numb. of Children	3.2	(2)	3.4	(2.2)	3.5	(2.2)
Black	0.41	(0.49)	0.4	(0.49)	0.45	(0.5)
Hispanic	0.11	(0.31)	0.059	(0.24)	0.13	(0.34)
Mom Works	0.15	(0.36)	0.19	(0.39)	0.19	(0.39)
Mom Works Full-time	0.076	(0.26)	0.064	(0.25)	0.06	(0.24)
Mom's Monthly Income	20	(61)	13	(39)	12	(36)
Monthly AFDC Payment	154	(89)	99	(65)	97	(56)
Monthly Household Income	193	(93)	130	(70)	126	(63)
Length of AFDC Spell (months)	41	(47)	51	(55)	47	(51)
Father Deceased	0.056	(0.23)	0.12	(0.33)	0.12	(0.33)
Father in the Home	0.19	(0.4)	0.32	(0.47)	0.26	(0.44)
Mom is a High School Grad	0.19	(0.39)	0.077	(0.27)	0.064	(0.24)
N (Households)	65234		2648		2407	

Note: Statistics are at the individual level in the top panel, and at the household level in the bottom panel. Roughly 5% of individuals are not matched to counties and therefore not used in this analysis. Mom's income, AFDC payment and household income are monthly and in nominal dollars

Table 2: RDs on County Head Start Funding 1966-1972

	(1) Combined	(2) 1966	(3) 1967	(4) 1968	(5) 1969	(6) 1970	(7) 1971	(8) 1972
A: Funding Dollars (Real 2016 $'s)								
RD_Estimate	453.91	94.87	393.46	428.26	136.11	795.00+	657.90+	175.38
	(301.72)	(249.44)	(393.73)	(370.81)	(374.25)	(458.09)	(399.05)	(285.29)
Robust P	0.181	0.674	0.411	0.287	0.763	0.119	0.122	0.626
Control Mean	539	442	525	564	639	669	753	532
Optimal BW	3.76	7.51	5.09	5.49	4.97	4.97	4.58	4.56
N	3315	452	308	332	302	302	277	277
B: Received Funding (0/1)								
RD_Estimate	-0.04	-0.12	0.09	-0.03	-0.11	0.01	0.06	-0.03
	(0.08)	(0.11)	(0.13)	(0.11)	(0.10)	(0.09)	(0.09)	(0.09)
Robust P	0.567	0.426	0.346	0.930	0.312	0.782	0.500	0.737
Control Mean	0.331	0.456	0.348	0.415	0.305	0.284	0.239	0.235
Optimal BW	6.43	6.09	4.78	6.91	6.64	8.30	6.24	6.20
N	6224	371	291	412	403	491	373	371
C: Dollars of Funding Conditional on Receipt								
RD_Estimate	1606.37*	437.79	469.82	1090.04	998.15	2519.61*	2238.79*	938.65
	(684.75)	(454.21)	(788.01)	(731.23)	(923.57)	(951.09)	(957.16)	(786.34)
Robust P	0.040	0.395	0.774	0.253	0.433	0.017	0.037	0.259
Control Mean	1808	1009	1467	1350	2066	2313	2830	2159
Optimal BW	4.07	9.66	4.63	5.92	5.45	5.10	4.83	5.76
N	1075	238	90	146	93	85	71	84

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Control mean is the estimate from the local linear regression to the left of the cutoff (the 301st poorest county). Alaska is not included. 1966-68 funding comes from the Community Action Program files, 1969-1980 funding comes from the Federal Outlay System Files. Funding dollars are scaled by number of 3-4 year olds in a county as estimated by SEER. The last column contains only counties who received positive funding.

Table 3: RDs on County Head Start Funding 1973-1980

	(1) 1973	(2) 1974	(3) 1975	(4) 1976	(5) 1977	(6) 1978	(7) 1979	(8) 1980
A: Funding Dollars (Real 2016 $'s)								
RD_Estimate	507.42	264.15	259.41	261.80	304.74	283.37	347.48	254.06
	(364.90)	(275.14)	(345.44)	(321.04)	(355.47)	(349.64)	(352.32)	(306.77)
Robust P	0.199	0.396	0.492	0.470	0.445	0.483	0.395	0.462
Control Mean	495	524	470	423	473	498	487	473
Optimal BW	5.15	4.50	4.88	4.85	4.83	4.85	5.04	5.01
Effective N	313	272	298	294	294	294	307	307
B: Received Funding (0/1)								
RD_Estimate	-0.04	-0.04	-0.09	-0.08	-0.09	-0.08	-0.07	-0.07
	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Robust P	0.668	0.716	0.445	0.457	0.393	0.470	0.462	0.477
Control Mean	0.266	0.266	0.283	0.247	0.244	0.246	0.237	0.244
Optimal BW	5.66	6.04	5.46	5.24	5.16	5.30	5.18	5.42
Effective N	346	369	330	319	314	321	315	329
C: Dollars of Funding Conditional on Receipt								
RD_Estimate	1805.99*	1122.79+	1448.21*	1481.49+	1882.66*	1730.67*	1974.52*	1500.09+
	(913.94)	(575.66)	(702.11)	(817.04)	(902.24)	(879.13)	(881.00)	(778.80)
Robust P	0.064	0.087	0.062	0.105	0.058	0.070	0.034	0.070
Control Mean	1979	1747	1638	1647	1799	1982	2085	2022
Optimal BW	6.27	5.58	5.70	6.03	5.72	5.69	5.75	5.71
Effective N	94	93	94	93	89	88	87	87

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Control mean is the estimate from the local linear regression to the left of the cutoff (the 301st poorest county). 1966-68 funding comes from the Community Action Program files, 1969-1980 funding comes from the Federal Outlay System Files. Funding dollars are scaled by number of 3-4 year olds in a county as estimated by SEER. The last column contains only counties who received positive funding.

Table 4: First Stage on Enrollment in Census and AFDC Data

	(1) All	(2) Non-White	(3) Single Mom	(4) on AFDC
	Narrow Measure			
	Nursery School			*Head Start*
A: Without covariates				
RD_Estimate	0.0268*	0.0503+	0.0301	0.058
	(0.0122)	(0.0260)	(0.0308)	(0.036)
Robust P	0.0665	0.0744	0.261	.38
B: With covariates				
RD_Estimate	0.0190*	0.0436+	0.0206	0.046
	(0.0079)	(0.0230)	(0.0277)	(0.035)
Robust P	0.0756	0.0479	0.467	.61
BW Mean	0.05	0.05	0.06	.03
	Broader Measure			
	Kindergarten or Nursery School			*Any School*
C: Without covariates				
RD_Estimate	0.0618*	0.153**	0.0827+	0.091*
	(0.0273)	(0.0475)	(0.0428)	(0.045)
Robust P	0.183	0.0107	0.124	.2
D: With covariates				
RD_Estimate	0.0340	0.147**	0.0774*	0.070
	(0.0217)	(0.0390)	(0.0367)	(0.044)
Robust P	0.395	0.0007	0.0977	.41
BW Mean	0.16	0.17	0.17	.06
N	421	330	378	357
Individuals	41000	14000	5600	1763

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty in Columns 1-3 and 12 percentage points for Column 4. Regressions are at the county level. The birth cohorts consist of all five year olds, the oldest three fourth of 4 year olds and the youngest one fourth of 6 year olds. Standard errors in parentheses, stars indicate conventional significance. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used. Covariates include: State fixed effects, white, high school degree, bachelors degree, married, years of education, mom's age, number in household, number of children and child's year of age and per capita county spending on means tested programs (via REIS).

Table 5: Reduced Form Effects on Mother of Head Start Aged Children in 1970 Census

	(1) (Placebo) Worked 1965	(2) Employed	(3) Hours	(4) Receive Welfare	(5) Welfare $'s	(6) In School
A: Single Mothers						
RD_Estimate	0.0167	-0.0694+	-3.193*	0.0065	54.27	-0.0198
	(0.0415)	(0.0407)	(1.550)	(0.0440)	(63.83)	(0.0150)
Robust P	0.770	0.0542	0.0261	0.840	0.457	0.266
BW Mean	0.34	0.45	15.52	0.36	441.0	0.04
County N	407	407	407	407	407	407
Individuals	4900	4900	4900	4900	4900	4900
B: Non-White Mothers						
RD_Estimate	0.0118	0.0655	0.967	-0.0297	-10.23	-0.0115
	(0.0387)	(0.0400)	(1.526)	(0.0237)	(34.63)	(0.0100)
Robust P	0.735	0.212	0.801	0.0763	0.373	0.220
BW Mean	0.26	0.41	13.34	0.17	201.1	0.02
County N	347	347	347	347	347	347
Individuals	12000	12000	12000	12000	12000	12000
C: All Mothers						
RD_Estimate	0.0198	0.0025	0.0969	-0.0017	3.551	0.0056
	(0.0159)	(0.0184)	(0.721)	(0.0069)	(9.694)	(0.0035)
Robust P	0.997	0.508	0.553	0.867	0.421	0.172
BW Mean	0.24	0.35	12.23	0.08	91.51	0.02
County N	421	421	421	421	421	421
Individuals	37000	37000	37000	37000	37000	37000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used. Covariates are included.

Table 6: Change in Composition of AFDC Recipients, Child born 1962-1963

	(1) Welfare Spell (Mnths)	(2) Spell Under 4mo	(3) Spell Under 12mo	(4) Spell Above Median	(5) Black	(6) Mom's Education	(7) Father in Home
A: Optimal Bandwidth							
RD_Estimate	9.913	-0.125^{+}	-0.188*	0.178^{+}	0.153	0.504	-0.213*
	(7.455)	(0.070)	(0.085)	(0.106)	(0.106)	(0.454)	(0.089)
Robust P-Val	.128	.0468	.0181	.0721	.144	.487	.0254
Bandwidth	5.9	5	6.1	5.5	8.2	7.1	7.2
Control Mean	32	.174	.362	.412	.51	7.63	.425
N	243	202	249	227	321	273	289
Individuals	3852	3284	4187	3616	5618	4912	5020
B: Bandwidth = 7							
RD_Estimate	8.126	-0.085	-0.179*	0.143	0.172	0.498	-0.212*
	(6.999)	(0.063)	(0.082)	(0.097)	(0.114)	(0.456)	(0.090)
Robust P-Val	.053	.054	.034	.07	.36	.7	.07
Control Mean	33.4	.154	.36	.433	.495	7.64	.423
N	279	279	279	279	279	270	279

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes all households with a child born 1962-1963, collapsed to the county level. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Control mean is the estimate from the local linear regression to the left of the cutoff (the 301st poorest county). Number of counties reported in "N", number of individuals underlying the counties is reported as "Individuals."

Table 7: Reduced Form Long-Run Effects on Mothers in 1980 Census - 10 Years After Head Start

	(1) Married	(2) Children Born	(3) Worked 1975	(4) Employed	(5) Hours	(6) Receive Welfare	(7) Welfare $'s	(8) In School	(9) Years of Ed
A: Single Mothers									
RD_Estimate		0.0602	-0.0404	-0.0436+	-2.588*	0.0181	131.2+	0.0229**	-0.0333
		(0.184)	(0.0319)	(0.0238)	(1.053)	(0.0219)	(76.73)	(0.0085)	(0.177)
Robust P		0.913	0.743	0.363	0.0648	0.216	0.0643	0.0547	0.872
BW Mean		5.37	0.50	0.54	19.29	0.36	858.6	0.03	9.94
N		420	420	420	420	420	420	420	420
Individuals		11000	11000	11000	11000	11000	11000	11000	11000
B: Non-White Mothers									
RD_Estimate	0.0836*	-0.349	-0.0325	-0.0705+	-3.267*	0.0246	37.70	0.0254**	-0.177
	(0.0381)	(0.274)	(0.0321)	(0.0410)	(1.329)	(0.0231)	(82.56)	(0.0071)	(0.288)
Robust P	0.0260	0.307	0.844	0.243	0.0243	0.294	0.413	0.0008	0.912
BW Mean	0.65	5.84	0.48	0.51	17.23	0.28	620.4	0.03	9.80
N	371	371	371	371	371	371	371	371	371
Individuals	17000	17000	17000	17000	17000	17000	17000	17000	17000
C: All Mothers									
RD_Estimate	0.0038	0.0941	-0.0010	-0.0030	-0.325	0.0059	17.11	0.0051	0.0548
	(0.0137)	(0.0979)	(0.0143)	(0.0145)	(0.559)	(0.0076)	(27.36)	(0.0035)	(0.0994)
Robust P	0.695	0.612	0.807	0.469	0.230	0.332	0.454	0.197	0.548
BW Mean	0.82	4.96	0.46	0.52	18.80	0.15	358.4	0.02	10.58
N	420	420	420	420	420	420	420	420	420
Individuals	54000	54000	54000	54000	54000	54000	54000	54000	54000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used. Covariates are included.

Table 8: Placebo Test of Kindergarten Enrollment in 1960 Census

	(1) All	(2) Non-White	(3) Not-Married
	Kindergarten		
A: Without covariates			
County level, without covariates			
RD_Estimate	-0.0133	-0.0132	0.0065
	(0.0197)	(0.0359)	(0.0421)
Robust P	0.203	0.695	0.868
B: With covariates			
RD_Estimate	-0.0120	0.0048	0.0030
	(0.0177)	(0.0325)	(0.0400)
Robust P	0.359	0.686	0.937
BW Mean	0.16	0.19	0.20
Counties	841	662	706
Individuals	55000	19000	5200

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes all children born between 1964 and the first quarter of 1965. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Bandwidth mean is the mean over the full bandwidth used. Number of counties reported in "N", number of individuals underlying the counties is reported as "Individuals."

Table 9: Placebo Tests of Maternal Outcomes in 1960 Census

	(1) Employed	(2) Hours	(3) Family Inc	(4) In School
A: Single Mothers				
RD_Estimate	0.0062	0.319	179.4	-0.0048
	(0.0306)	(1.094)	(136.9)	(0.0093)
Robust P	0.768	0.377	0.350	0.835
BW Mean	0.39	19.09	1932.0	0.03
N	413	413	413	413
Individuals	4200	4200	4200	4200
B: Non-White Mothers				
RD_Estimate	-0.0112	-0.376	139.5	-0.0138
	(0.0292)	(1.038)	(142.9)	(0.0090)
Robust P	0.493	0.446	0.615	0.151
BW Mean	0.27	14.52	2132.2	0.02
N	370	370	370	370
Individuals	26000	26000	26000	26000
C: All Mothers				
RD_Estimate	0.0082	0.323	-48.36	-0.0013
	(0.0128)	(0.429)	(61.57)	(0.0052)
Robust P	0.573	0.482	0.341	0.667
BW Mean	0.23	14.43	3265.5	0.02
N	421	421	421	421
Individuals	79000	79000	79000	79000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes all children born between 1964 and the first quarter of 1965. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Bandwidth mean is the mean over the full bandwidth used. Number of counties reported in "N", number of individuals underlying the counties is reported as "Individuals."

Appendix

A The Impact of Primary School on AFDC Mothers

To compare the findings on maternal labor supply I adapt the strategy used in Gelbach (2002) and Fitzpatrick (2012) to my sample of AFDC recipients. Namely, I estimate a regression discontinuity using month of birth and school age eligibility cutoff to identify the discontinuous increase in kindergarten enrollment for children born in December 1962 versus January 1963.[44]

Similarly to my strategy with Head Start, I estimate:

$$Y_{i,m} = b_0 + b_1(birthmonth_{i,m} - Dec1962) + \alpha G_m + G_m b_2(birthmonth_{i,m} - Dec1962) + \nu_{i.m} \quad (3)$$

where in the first stage, $Y_{i,m}$ measures an individual i born in month and year m. In the reduced form on mothers, $Y_{i,m}$ is an outcome for a woman whose child was born in month and year m.

Table A12 presents the increase in kindergarten enrollment for children at the birth month discontinuity. Children born in December of 1962 are 15 percentage points more likely to be enrolled in Kindergarten than those born a month later. This is despite the fact that in 1967 21 states did not yet offer funding for school districts to offer kindergarten. If I look only at the states were kindergarten was widely available at this time (Columns 2-5), the impact rises to 20-30 percentage points (depending on bandwidth). The mean kindergarten enrollment for children born a month after the cutoff is roughly 5%, making this an up to 600% increase.

Figure A10a offers further proof that this cutoff was highly predictive of kindergarten enrollment. It plots the estimated discontinuity if we use a different age cutoff (one every

[44] In 1967 most of the country used the calendar year to define school start age: children should turn 5 in the year they began kindergarten Elder and Lubotsky (2009).

two-months). The coefficient at zero is the one in Column 2 of Table A12. Only one other alternative cutoff within 7 months on either side of this is significantly different from zero (and is much smaller in absolute value).

The first panel of Table A13 shows the reduced form effects of the school age cutoff on the mothers of the children. Even though the impact on enrollment for this is much stronger than the one I find for Head Start, there are no impacts on their mothers. The 30 percentage point increase that a child is enrolled in Kindergarten yields no change in the composition of welfare families. If mothers were leaving welfare after their child entered kindergarten, we should see some change in the composition of welfare spells. There is also no change in the composition by maternal education or presence of father. This further suggests that Head Start is a far different intervention than kindergarten or other childcare and schooling programs. Additionally, since no change in the composition of AFDC mothers makes it unlikely that becoming eligible for school caused mothers to return to work, this population may be one whose labor supply is not very responsive to the availability of free childcare.

The population I've just considered has very little overlap with the sample I use in the RD for Head Start, this is because the counties near the cutoff are located in states without kindergarten programs. Thus I repeat the process for entry into first grade using December 1961 as the birth month cutoff. This discontinuity yields a similarly large increase in school attendance for children (Figure A10b). To get closer to the sample that inform my Head Start estimates, I restrict the estimates to counties that fall within ten percentage points of the 300th poor county cutoff. These results are shown in Panel C of A13. Here again there are no changes in the composition of welfare recipients due to school eligibility. These results show that Head Start acts differently upon mothers as compared to regular primary school, and that the differences aren't simply due to the sample of very disadvantaged counties.

B Data Appendix

B.1 Individual Level Data

I showed in Section 4 that the density of counties across the cutoff is continuous, and thus I perform all my analyses at the county-level. Interestingly though, if we perform the same test using the density of the running variable at the individual level, we see a different picture (Figure A8a). The density is actually discontinuous, there are more people just under the cutoff than just over. This is because there is a population outlier county that was just a fraction of a percent less poor than the cutoff. Cameron County, Texas has a population that is more than twice as large as any other county within five percentage points of the cutoff. As mentioned earlier, it's implausible that this density difference is due to actual manipulation by anyone, but it does tell us that coincidentally, there's a single county that may have an unduly large impact on the estimation. In Figure A8b I repeat the individual-level density after dropping Cameron County, and the density is again smooth through the cutoff.f

The individual level data passes the McCrary test after one county is dropped, but this result emphasizes a broader pattern in the data. Because the population sizes of the counties have such high variation, and individual characteristics are correlated within county, the distribution of individuals over the running variable is likely to be discontinuous while the distribution of counties is not. For this reason I estimate the RD regressions giving each county equal weight. However my main results are quite similar using individual or population-weighted estimates (shown in Tables A7 and A8). I also control flexibly for county population size in my main specifications.[45]

[45] I include indicators for quartiles of rural-ness and an indicator for counties that are in the top percentile of population in the sample.

B.2 Verifying the Running Variable

The running variable, 1960 county poverty rate, was created from the 1960 Putnam file. The Bureau of the Census created the Putnam File for OEO to identify the extent and patterns of poverty at the State and county/city levels.[46]

> *To establish this definition and pattern of poverty, the Bureau of the Census took the first step and created the PUTNAM FILE using data from the Census of 1960... From this basic 1960 Putnam File, OEO created the 1960 PUTNAM AGGREGATE PRINT FILE using OEO definitions of poverty and various statistical aggregating techniques... They were the primary source for OEO's Community Profiles and for its reports to Congress on poverty levels by state, county, and congressional District. OEO also used these files to justify the allocation of funds... (NARA Documentation of the Putnam File)*

So, while we can be relatively sure that this file was how OEO measured poverty, the file contains multiple measures of poverty (individual, family and child poverty are all reasonable options). The Putnam file is housed by the National Archives and Records Administration (Office of Economic Opportunity, 1975). The current publicly available file is in an obsolete EBCDIC plain text format, making it labor intensive to read. Martha Bailey and Nic Duquette did this translation and provided me with a copy of their cleaned file. Separately, Jens Ludwig shared the SAS file he obtained from NARA in 2002. I verified that these two were identical.

Jens Ludwig's 2002 files contained more extensive documentation than NARA currently provides. It appears that NARA disposed of the Putnam Aggregate file in 2003 "because of unresolvable problems with the preservation media." At this time they also appeared to remove the documentation that explains how to convert the Putnam File to the Putnam Aggregate File. I am very fortunate to have access to the old version. I followed the written

[46]The idea of a "poverty threshold" dates to 1964.

instructions to derive the Putnam Aggregate File. For example, the Putnam Aggregate variables I use to construct the poverty rate are: *Variable 5* (Persons), and *Variable 6* (Persons below poverty cutoff). *Variable 5* is defined as *Variable 18* (White Persons) + *Variable 19* (Non-white persons). Variable 18 is defined as *"Putnam File, Table IIA, IIB (# families * family size) + (# persons in families of 16 or more) + (# primary and secondary individuals in households)."* Where the tables are part of the documentation that defines the meaning of the variables in the Putnam file (the variables are numbered 1-465).

While completing this process I discovered that the running variable used in Ludwig and Miller (2007) was wrong. During data cleaning they multiplied the *"persons in families of 16 or more"* by 16, so the denominator for the poverty rate was always off by a multiple of 16. This had a very small impact on the overall poverty rates for the counties (the correlation between the two series is very high), but it did result in counties poverty ranking changing, making 8 counties switch from one side of the cutoff to the other (the counties that were previously ranked 295, 296, 297 and 299 no longer fall within the 300 poorest). In an RD setting moving counties from one side to the other is potentially very meaningful.

Although *"persons below poverty cutoff"* over *"persons"* is the most obvious choice for the OEO to have ranked counties, I also compared other potential poverty measures. Namely: the proportion of families in poverty and the proportion of families earning less than $3000 per year (which was a common measure of disadvantage before the poverty threshold was established). In addition I created an alternate estimate of individual poverty from the Putnam file that would have been the intuitive choice had their not been documentation on OEO's aggregation method.[47]

Table A1 reports how these poverty rates differ. *LM07* is the measure from Ludwig and Miller (2007) with the error, *Correction* is my preferred measure, *Alternate* is the intuitive aggregation. All are highly correlated, but between 8-120 counties change cutoff sides depending on the comparison. In Tables not included here I compare how predictive these

[47]The basic Putnam file contains hundreds of variables and the brief description provided does not always make it clear exactly who is included.

variables are of Head Start funding (available by request) and Census nursery school enrollment. My preferred measure, although chosen initially because it made the most sense, is also the most predictive of a discontinuity.

B.3 Over and Under-counting of Head Start in Surveys

One likely reason for the widely different estimated increases in Head Start enrollment between the 1970 Census and the 1988 NELS is due to reporting errors. In this section I compare the implied attendance levels in these surveys to administrative records of national Head Start attendance. This exercise shows that too few children report being in nursery school in the 1970 Census and too many parents report their child attended Head Start in the NELS. Laughlin and Davis (2011) have previously identified issues of Head Start miscount in the ACS, CPS and SIPP, but not to my knowledge this has not been previously documented in the Census or NELS.

To get approximate Head Start attendance in the 1970 Census I look at all children aged 3-5 who have family incomes within 125% of the poverty threshold and report attending any type of nursery school. This number amounts to 114,000 near-poor children attending nursery school in April 1970.[48] Head Start reported that they had 250,000 full-year enrollees in 1970. This means that over half of attendees are not accounted for, and this measure includes nursery schools that are not Head Start (although this was rare).[49]

In contrast, the NELS reports that 15% of the (weighted) NELS sample reported attending Head Start. If we assume that the 1988 8th grade cohort included about 3.3 million students, this implies that 495,000 member of that cohort attended Head Start. Head Start enrollment for each year 1978-1980 was approximately 350,000 students. These numbers are a less straight forward comparison because this cohort could have attended Head Start for two years, but at the same time there would have been an older or younger cohort attending

[48] I also use 150% of the poverty line, which increases the count somewhat but still falls short of the administrative number.

[49] By 1980 there are too many other types of preschool available and this approach is no longer helpful.

alongside them. If we look at these numbers naively we estimate that there is roughly 40% over reporting.

If we are more careful we can make the following two assumptions: first, assume that all of these children attended in the year before they started primary school (e.g. some spent multiple years but none dropped out) and second, that 70% of children of children in a center are in their final year.[50] These two assumptions mean imply that only 70% of the 350,000 slots from the year before this cohort began primary school would be available to them (245,000), and that their reported enrollment is double what it should be.

B.4 Matching Counties in AFDC 67 Survey

Andrew Goodman Bacon matched 2,429 counties in the AFDC data to FIPS codes by assuming that the numerical code assigned to each county in the data corresponded to the county's alphabetical ranking. Counties too small to appear in the survey were skipped, leaving telltale gaps in the numerical code. It also appeared that a "space" in a county name corresponded to the last alphabetical ordering instead of the first.[51] He confirmed these matches using administrative counts of AFDC caseloads.

This method does not work for Alaska, Connecticut, Hawaii, Illinois, South Dakota, Massachusetts, New York, or Virginia whose county codes clearly follow different patterns. Through careful scrutiny I have identified all the counties except for those in Hawaii, Virginia and Massachusetts (and half of Connecticut).

Alaska has split some of its counties into two, but the first digit of the code corresponds to the counties as elsewhere listed in 1968. However these counties do not map easily to modern Alaskan counties and I therefore drop them.

Connecticut has only 8 counties which are not listed in any particular order. However the more populous four counties are different enough in caseload that assigning them in caseload

[50] I believe that this is an overestimate, Ways and Means Committee (2008) reports that 24% of attendees in 1980 were 3, 55% were 4 and 21% were 5.

[51] E.G. El Paso, TX comes after Ellis County instead of before.

order seems reasonable. I assign these four and drop the smaller four.

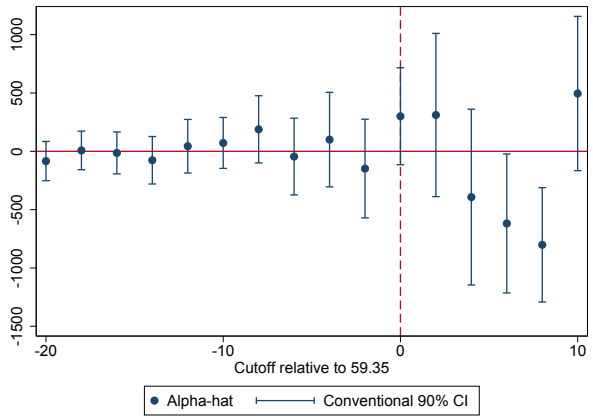
Illinois splits Chicago/Cook county into 20 counties that occupy the "200s" of the county codes. The remaining counties are in alphabetical order (but skip Cook County).

New York lists 73 counties, while the state only has 58. Those in the 100s are all parts of New York City, the first digit of the remaining codes seems to be some kind of regional indicator. If the first digit is removed than the remaining two digits matches the alphabetical county list. There are five sub-counties that are reported in other lists of counties from the same period (Auburn, Binghamton, Jamestown, Poughkeepsie, Union). It seems like these were listed in the last five slots, in alphabetical order. I add these back to their parent counties.

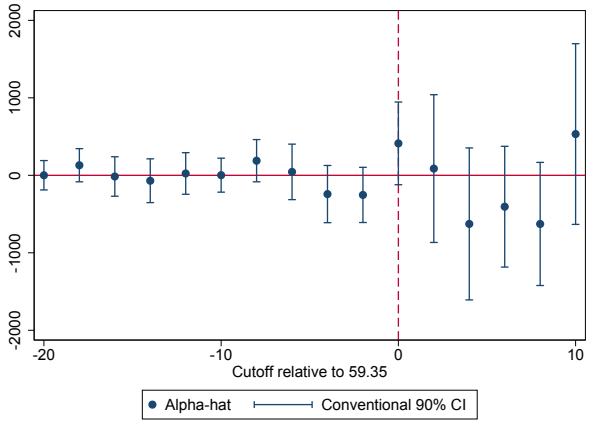
In South Dakota, the largest county, Minnehaha appears to be listed first. The remaining counties except for three are listed alphabetically. Those out of order three are Shannon, Washabaugh and Todd. These counties are also listed separately in the 1968 caseload data. They are listed in a different order, but since they vary enough in size it's easy to guess which is which.

Figure A1: Varying the Cutoff: Real Funding per Child (Selected Years)

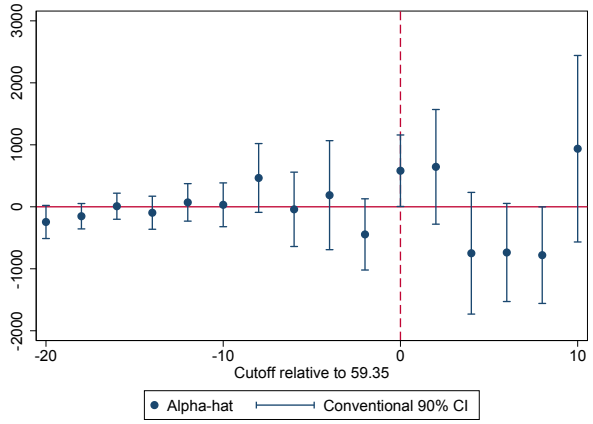
(a) 1966-1980 Combined



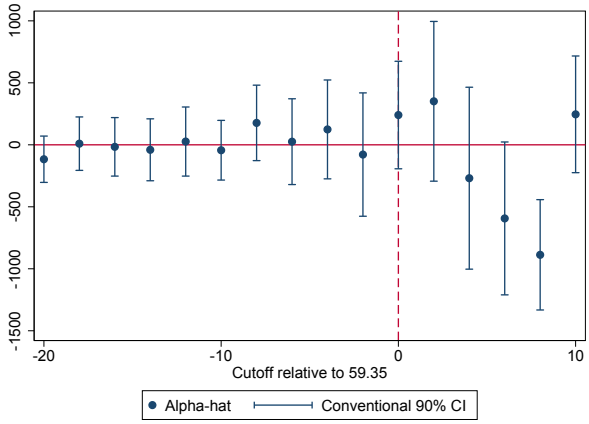
(b) 1967



(c) 1970



(d) 1980

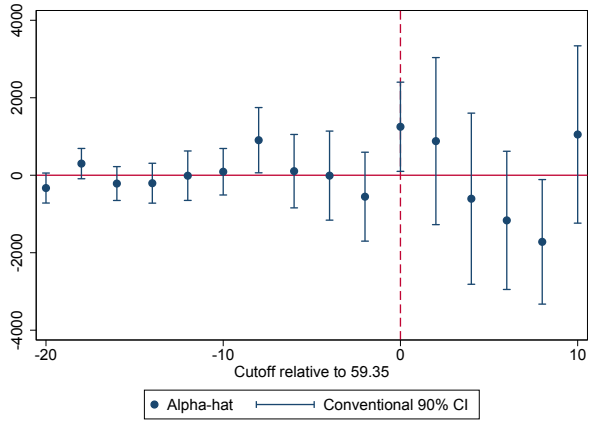


Note: I move the cutoff 2 percentage points of county poverty and re-estimate the RD. The point estimate is graphed along with 90% conventional confidence intervals. The outcomes in this figure are all overall funding per child (including counties with 0 dollars of funding). Bandwidth fixed at 10. Data comes from CAP and FOSF.

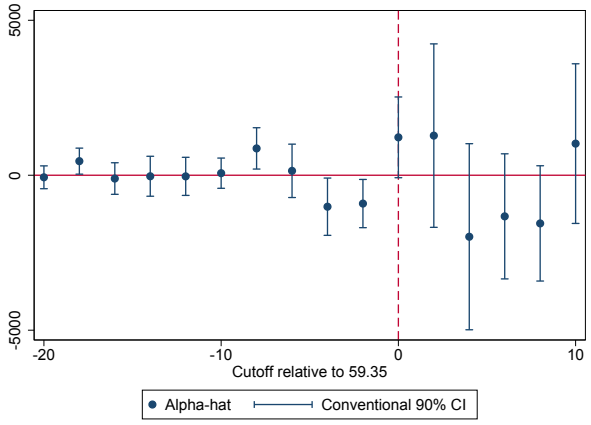
C Appendix Figures

Figure A2: Varying the Cutoff: Intensive Funding (Selected Years)

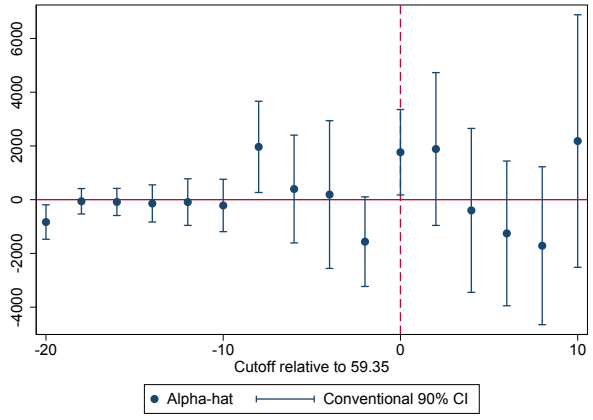
(a) 1966-1980 Combined



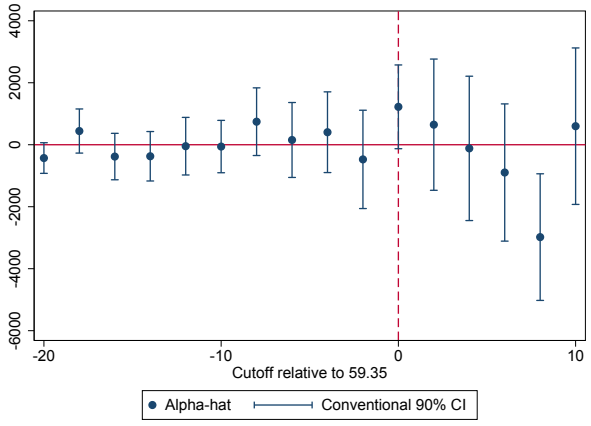
(b) 1967



(c) 1970



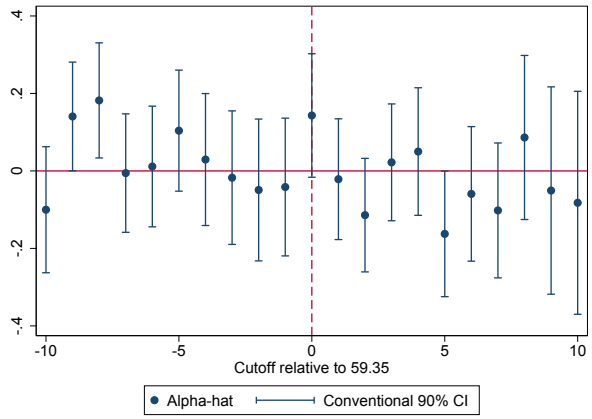
(d) 1980



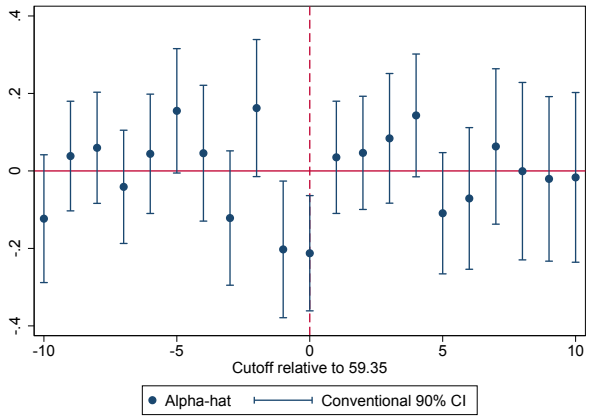
Note: Intensive funding refers to real funding dollars among the subset of counties that receive any funding. I move the cutoff 2 percentage points of county poverty and re-estimate the RD. The point estimate is graphed along with 90% conventional confidence intervals. Bandwidth fixed at 10. Data comes from CAP and FOSF

Figure A3: Varying the Cutoff: Composition of Welfare Recipients in 1967

(a) Spell Length Above Median



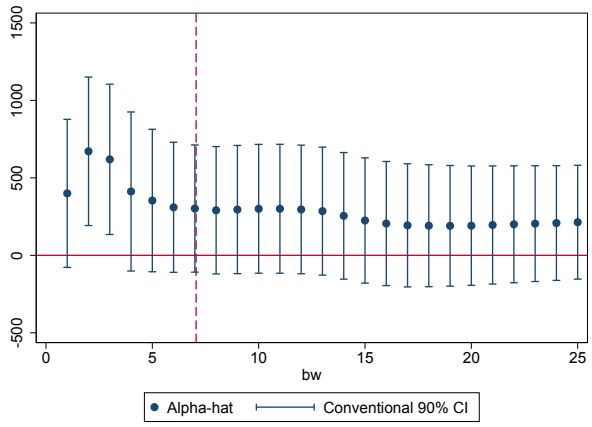
(b) Father Present in Home



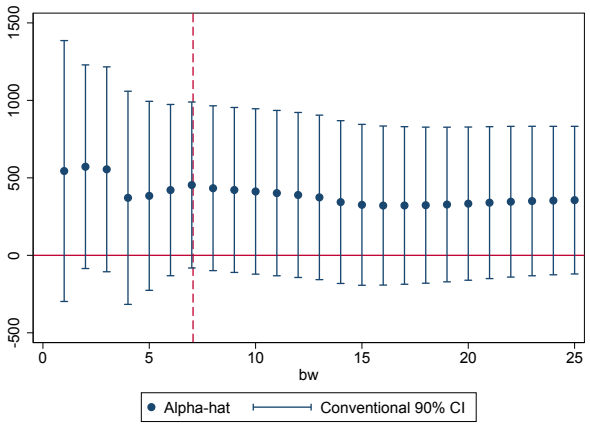
Note: I move the cutoff 1 percentage points of county poverty and re-estimate the RD. The point estimate is graphed along with 95% conventional confidence intervals. The sample used in these figures is households in the 1967 AFDC data which contain a child born 1962-1963. Bandwidth is fixed at 7.

Figure A4: Varying the Bandwidth: County Funding Per Child (Selected Years)

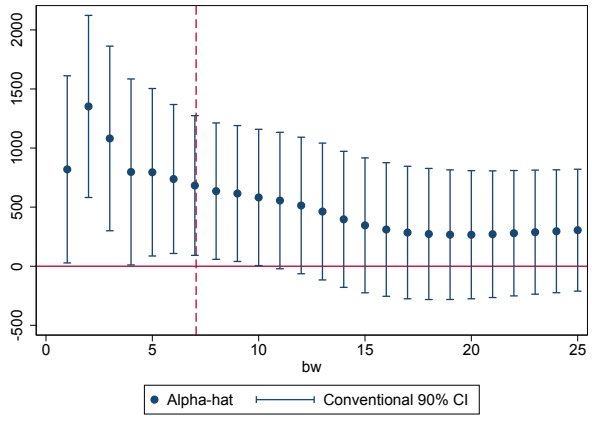
(a) 1966-1980 Combined



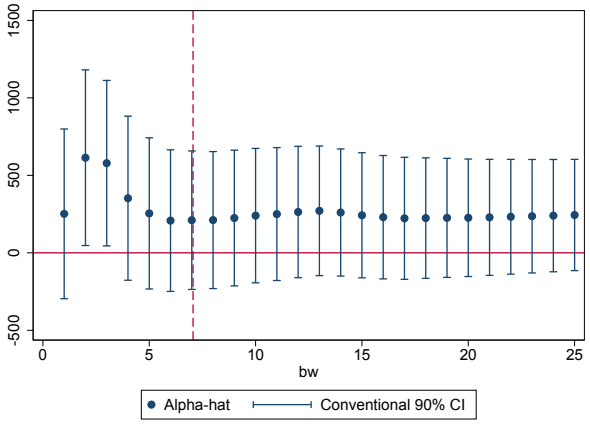
(b) 1967



(c) 1970

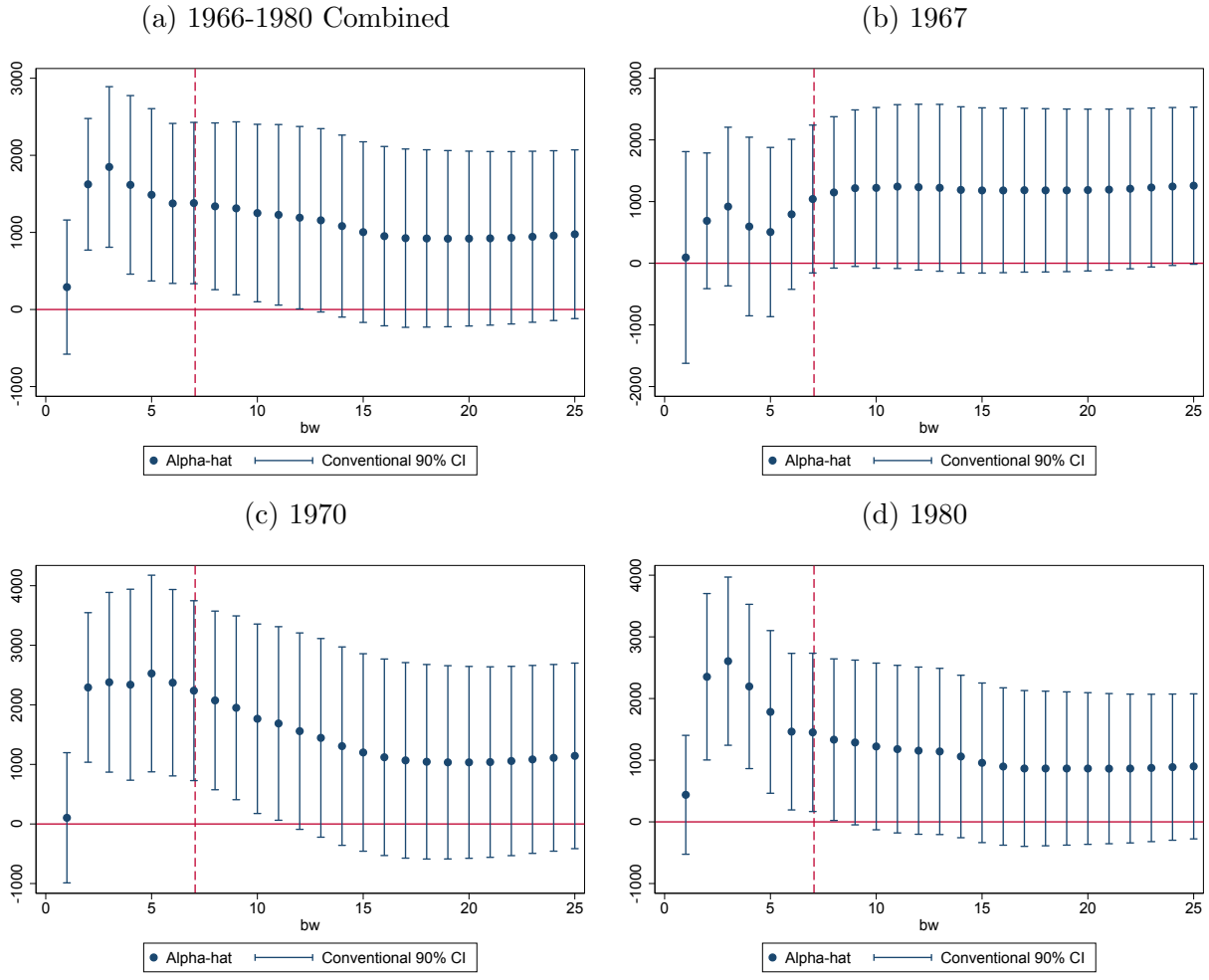


(d) 1980



Note: I re-estimate the RD coefficient for each integer bandwidth between 1 and 25. The point estimate is graphed along with 90% conventional confidence intervals. The outcomes in this figure are all overall funding per 3-4 year old child (including counties with 0 dollars of funding). Data comes from CAP and FOSF.

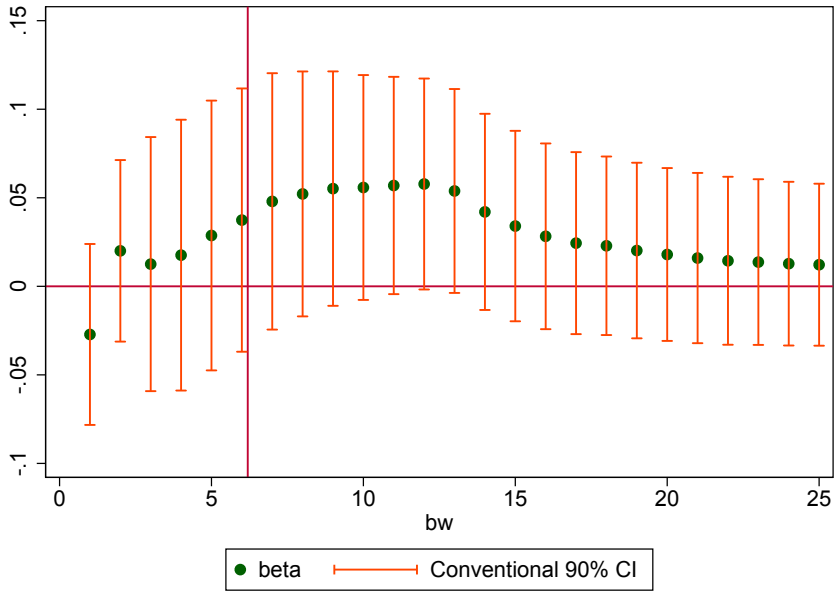
Figure A5: Varying the Bandwidth: Intensive Funding (Selected Years)



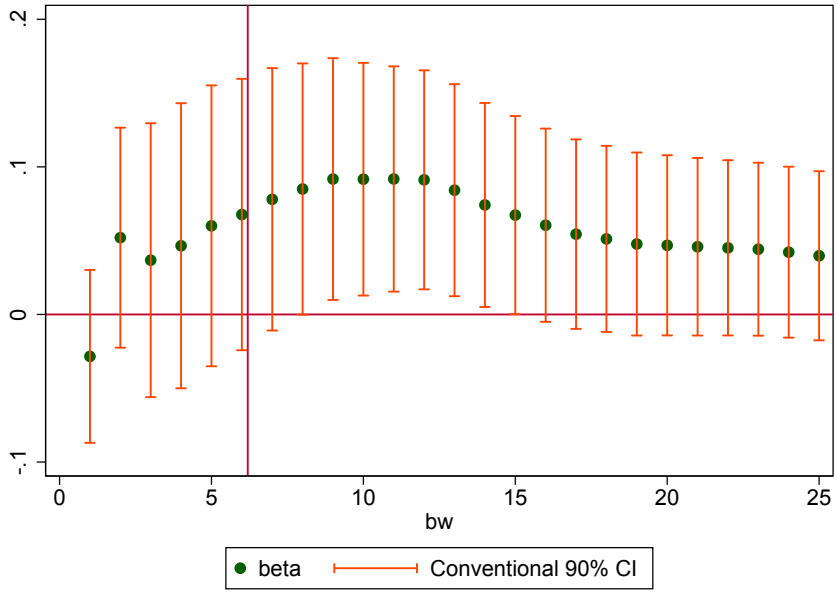
Note: Intensive funding refers to real funding dollars per child among the subset of counties that receive any funding. I re-estimate the RD coefficient for each integer bandwidth between 1 and 25. The point estimate is graphed along with 90% conventional confidence intervals. Data comes from CAP and FOSF.

Figure A6: Estimates of Discontinuity in Attendance and Confidence Intervals for different Bandwidths (born 1962-1963)

(a) Enrolled In Head Start



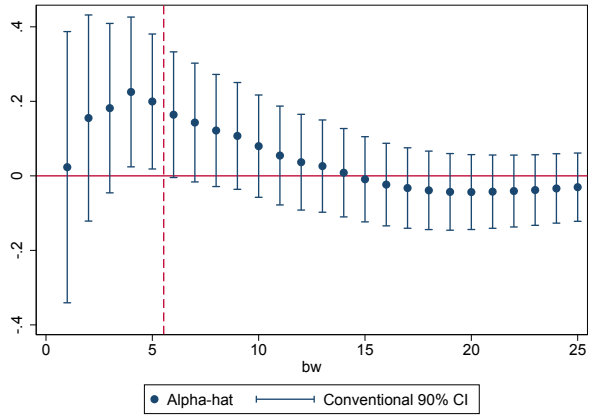
(b) Enrolled in Any School



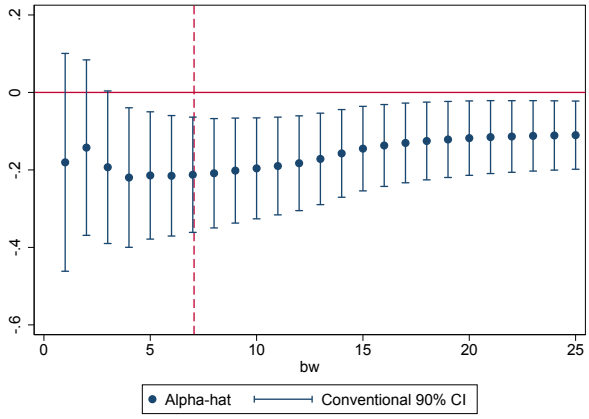
Note: These figure plots the estimated jump in attendance ($\hat{\alpha}$) for AFDC recipient children born 1962-63 (Aged 4-5) at the discontinuity for different choices of bandwidth. Bandwidth is measured in percentage points of 1960 county poverty and ranges from within one percentage point of the cutoff, to within 25 percentage points of the cutoff. It also shows the conventional 90% confidence interval for each estimate. The vertical line is the MSE optimal bandwidth choice suggested by CCT (2014).

Figure A7: Varying the Bandwidth: Composition of Welfare Recipients in 1967

(a) Spell Length Above Median



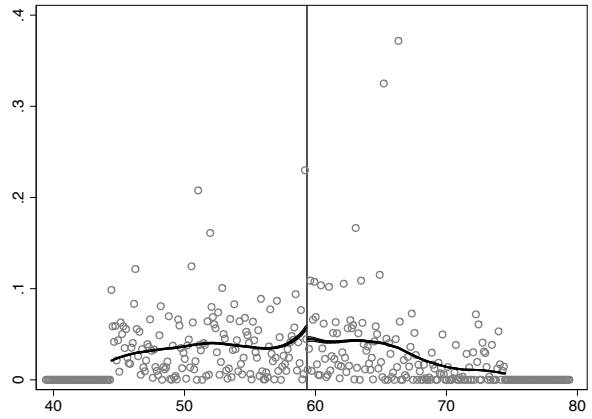
(b) Father Present in Home



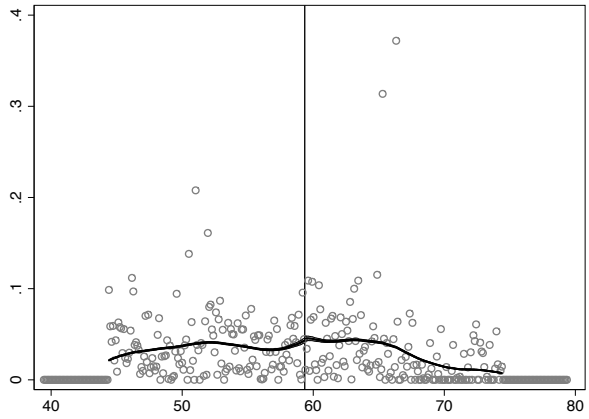
Note: I re-estimate the RD coefficient for each integer bandwidth between 1 and 25. The point estimate is graphed along with 90% conventional confidence intervals.

Figure A8: McCrary Test of Density of the Running Variable: Individual Level

(a) Individual Level



(b) Individual Level, Dropping Cameron County

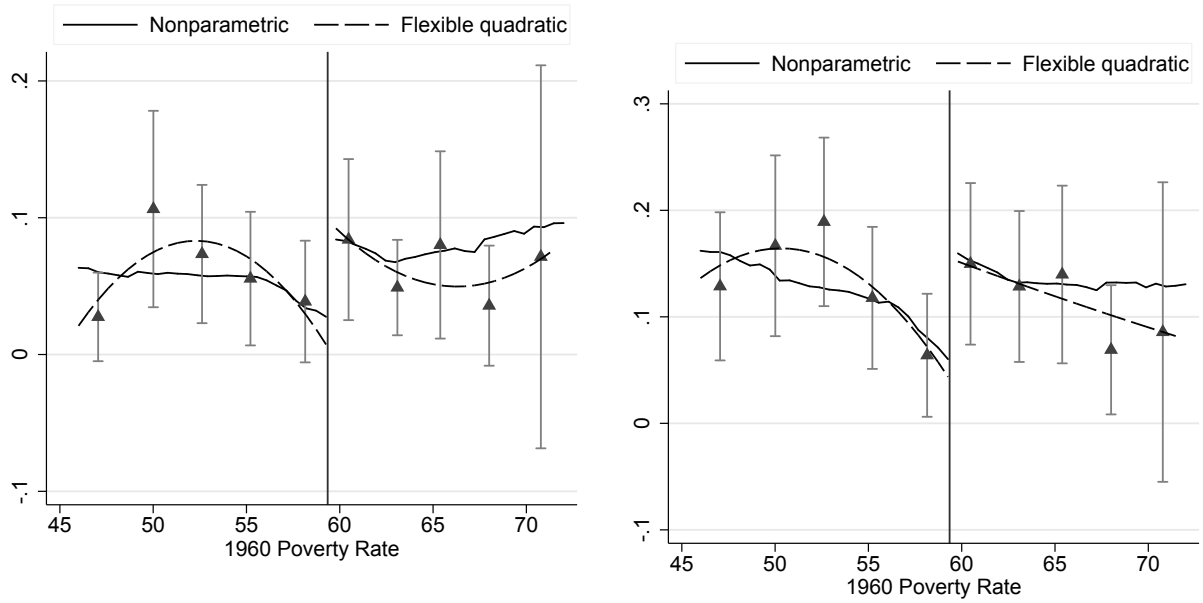


Note: Figures created using Stata add on "DCdensity" written by Justin McCrary. Figures use individual level data for AFDC recipients in 1967, matched to the 1960 county poverty rate variable from the Census Putnam File.

Figure A9: Graphical Depiction of Discontinuity in Attendance (born 1962-1963)

(a) Enrolled In Head Start

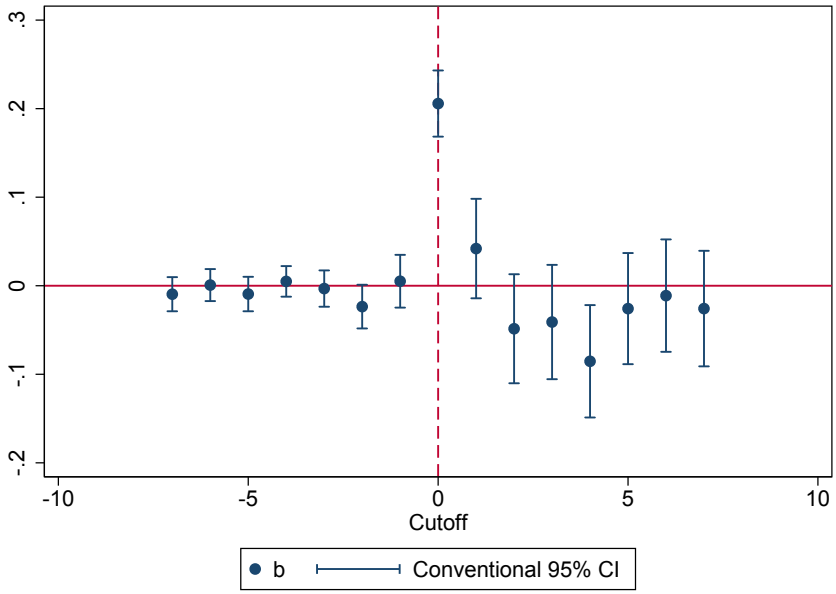
(b) Enrolled in Any School



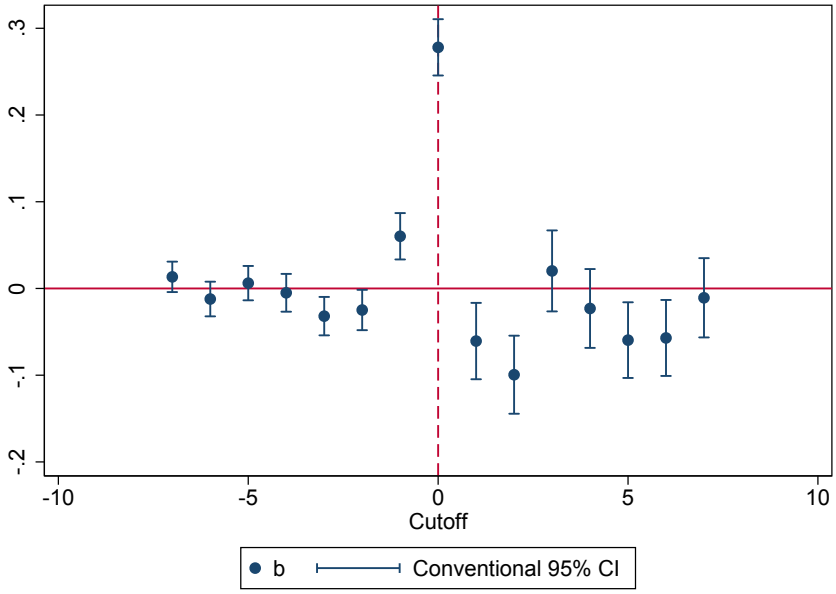
Note: These figures mimic the estimated discontinuities in Panels C and E of Table A4 (Column 1). They plot the local polynomials relating the outcome variable to the running variable (bandwidth=12). Bin means and confidence intervals are also plotted. Data come from children born 1962-1963 in the 1967 AFDC recipient sample. Data is collapsed to the county level using survey weights before estimation.

Figure A10: First Stage of Birth Month on School Enrollment

(a) Varying the cutoff point for Kindergarten age eligibility



(b) Varying the cutoff point for First grade age eligibility



Note: Coefficients for an RD at noted cutoff (bw=4 months). Bars are 95% confidence intervals. Children get older towards the right. First figure uses cutoff of December 1962, second uses December 1961. Only states that fully funded kindergarten programs before 1966 are included in the first figure..

D Appendix Tables

Table A1: Count of Switching Counties and Rank Order Correlations

	1960 Poverty Rate Measure				
	LM07	Correction	Alternate	Family Pov.	Under $3000
LM07		8	38	62	116
Correction	1.0000		40	64	120
Alternate	0.9988	0.9988		60	112
Family Pov.	0.9949	0.9948	0.9957		98
Under $3000	0.9715	0.9712	0.9711	0.9750	

Note: Top right numbers refer to the count of counties that differ on side of the cutoff. By construction these are even numbers, since counties on one side must switch with counties on the other. Bottom left numbers are correlations in the rank order of counties. Data come from Putnam file. "LM07" refers to the running variable using in Ludwig and Miller (2007), "Correction" refers to the corrected version used in this paper. See text for further detail.

Table A2: Decennial Census Summary Statistics by County and 1960 Poverty Rate

	All Counties		1960 Poverty 49.35-59.35%		1960 Poverty 59.35-69.35%	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
1970						
Family Poverty	0.11	(0.07)	0.27	(0.047)	0.35	(0.053)
Non-White	0.12	(0.12)	0.23	(0.17)	0.33	(0.22)
Households with Female Head	0.11	(0.034)	0.12	(0.028)	0.14	(0.028)
Under 18	0.34	(0.036)	0.36	(0.045)	0.39	(0.04)
Completed HS (aged 25+)	0.52	(0.11)	0.32	(0.062)	0.29	(0.05)
BA Degree	0.11	(0.047)	0.056	(0.025)	0.055	(0.02)
Female Share of Labor Force	0.38	(0.032)	0.37	(0.043)	0.37	(0.041)
Total Labor Force Participation	0.71	(0.064)	0.65	(0.085)	0.64	(0.074)
Median Income	9661	(2073)	5887	(741)	5142	(718)
Average Population	64476	(227973)	18874	(15766)	17566	(17453)
Number of Counties	3140		346		223	
1980						
Family Poverty	0.096	(0.031)	0.18	(0.041)	0.23	(0.049)
Non-White	0.17	(0.097)	0.23	(0.16)	0.32	(0.21)
Households with Female Head	0.14	(0.033)	0.14	(0.036)	0.16	(0.044)
Under 18	0.28	(0.022)	0.31	(0.035)	0.33	(0.038)
Completed HS (aged 25+)	0.66	(0.069)	0.47	(0.07)	0.43	(0.057)
BA Degree	0.16	(0.041)	0.089	(0.035)	0.085	(0.027)
Female Share of Labor Force	0.43	(0.018)	0.41	(0.039)	0.41	(0.04)
Total Labor Force Participation	0.76	(0.038)	0.7	(0.075)	0.69	(0.063)
Median Income	16973	(2371)	12098	(1759)	10951	(1605)
Average Population	213119	(4095776)	22386	(19896)	20517	(23873)
Number of Counties	3189		347		223	

Note: Statistics are population weighted means of county level data from the 1972 and 1983 City and County Data Books (which is derived from the Decennial Census). Average County population is a raw mean, and not weighted by population.

Table A3: CAP data: Non Head-Start funding

	(1) 1966	(2) 1967	(3) 1968	(4) 1969	(5) 1970	(6) 1971
			Real PC CAP Admin Funds			
RD_Estimate	3864.8	7722.2^{+}	18985.9	18063.6	4450.4	14837.6
	(3097.1)	(4295.4)	(15534.8)	(16464.0)	(14188.6)	(13399.5)
Rob p-val	0.261	0.0981	0.350	0.437	0.959	0.541
Base	3779.3	4436.6	31007.1	28766.5	37030.9	32418.4
BW	7.879	7.708	9.260	7.786	5.529	5.197
N	469	463	531	465	334	316
			Real PC Legal Services Funds			
RD_Estimate	-668.4	1035.0	-346.5	293.0	-485.7	469.1
	(820.7)	(1239.5)	(1546.6)	(487.7)	(476.1)	(417.5)
Rob p-val	0.541	0.461	0.796	0.265	0.546	0.127
Base	456.8	417.6	918.7	455.3	948.9	987.0
BW	9.504	8.817	11.36	3.649	4.853	4.352
N	544	509	618	215	294	263
			Real PC CAP Health Funds			
RD_Estimate	359.0	-519.4	113.7	0.0624	-2032.1	-1002.8
	(1132.2)	(1537.8)	(2071.9)	(0.0742)	(5777.8)	(4099.2)
Rob p-val	0.892	0.547	0.841	0.334	0.602	0.615
Base	721.0	1106.8	3548.3	0.537	9840.2	8799.1
BW	6.360	5.428	5.382	6.100	4.507	3.954
N	383	330	328	371	273	235
			Per Capita Senior Funding			
RD_Estimate	20.77	-224.0	-102.4	-400.5	-92.37	44.42
	(21.15)	(197.5)	(276.6)	(382.6)	(116.0)	(34.71)
Rob p-val	0.404	0.252	0.887	0.264	0.724	0.179
Base	24.36	57.86	142.4	344.0	235.7	94.44
BW	10.70	14.61	7.871	9.806	7.170	2.971
N	593	784	468	562	434	179

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Note: Data come from Bailey and Goodman-Bacon (2015) (https://www.aeaweb.org/aer/data/10503/20120070_data.zip). The underlying source is the same CAP grant data described in Section 3.3. Community Health Center funding is not included because is not present in enough counties near the discontinuity. There is no evidence of a discontinuity in the non-Head Start programs. The coefficients on administrative funding (top panel) are all positive, and significant at the 10% level in 1967. This could suggest that some of the Head Start administration was done by the broader CAP organization.

Table A4: First Stage on Head Start Attendance, Cohorts born 1961-1964 (Aged 3-6)

	(1) 1962-1963 Aged 4-5	(2) 1961 Age 6	(3) 1962 Age 5	(4) 1963 Age 4	(5) 1964 Age 3
A: MSE Optimal Bandwidth					
RD_Estimate	0.022	0.026	0.027	0.043	-0.010
	(0.046)	(0.017)	(0.061)	(0.057)	(0.021)
Robust P	.963	.231	.96	.665	.647
Bandwidth	4.6	8.2	5.2	6.3	6.4
Control Mean	.0437	-.00028	.0473	.0271	.00555
N	294	274	170	194	193
B: Bandwidth=12					
RD_Estimate	0.058	0.029*	0.060	0.054	-0.013
	(0.036)	(0.015)	(0.048)	(0.045)	(0.015)
Robust P	.38	.23	.44	.54	.54
C: With Covariates, Bandwidth=12					
RD_Estimate	0.046	0.027+	0.048	0.040	-0.011
	(0.035)	(0.015)	(0.046)	(0.043)	(0.016)
Robust P	.61	.27	.68	.73	.63
Control Mean	.0279	-.00704	.0342	.0215	.0227
D: Outcome: Enrolled in Any Kind of School, Bandwidth=12					
RD_Estimate	0.091*	-0.022	0.079	0.101+	-0.024
	(0.045)	(0.058)	(0.065)	(0.052)	(0.017)
Robust P	.2	.16	.64	.093	.35
Control Mean	.0576	.888	.0862	.0278	.0242
N (Counties)	358	357	334	324	309
N (Individuals)	1763	1031	899	864	797

Note: + p<0.10, * p<0.05, ** p<0.01. Years at the top of the column denote the birth cohort, approximate age in years in 1967 is below. Regressions are at the county level, with means of attendance among the cohort(s) indicated. Survey weights are used in aggregation to county level. Column one includes the two cohorts where Head Start is most common, 1962 and 1963. Standard errors in parentheses, stars indicate conventional significance. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Control mean is the estimate from the local linear regression to the left of the cutoff (the 301st poorest county). Covariates include: indicators for white, Hispanic, high school completion and if the father of the household is in the home. Also the number of children under 6, aged 6-12 and 13-17 in the household as well as the length of the current welfare spell in months and an indicator for the county being among the most populous in the sample.

Table A5: Reduced Form Effects on Mothers of Head Start Aged Children in 1970 Census Without Covariates

	(1) Worked 1965	(2) Employed	(3) Hours	(4) Receive Welfare	(5) Welfare $'s	(6) In School
A: Single Mothers						
RD_Estimate	-0.0038	-0.113**	-4.583**	0.0368	120.7	-0.0183
	(0.0480)	(0.0432)	(1.642)	(0.0524)	(83.93)	(0.0172)
Robust P	0.926	0.0075	0.0063	0.530	0.152	0.207
BW Mean	0.34	0.45	15.52	0.36	441.0	0.04
County N	407	407	407	407	407	407
Individuals	4900	4900	4900	4900	4900	4900
B: Non-White Mothers						
RD_Estimate	0.0118	0.0655	0.967	-0.0297	-10.23	-0.0115
	(0.0387)	(0.0400)	(1.526)	(0.0237)	(34.63)	(0.0100)
Robust P	0.735	0.212	0.801	0.0763	0.373	0.220
BW Mean	0.26	0.41	13.34	0.17	201.1	0.02
County N	347	347	347	347	347	347
Individuals	12000	12000	12000	12000	12000	12000
C: All Mothers						
RD_Estimate	-0.00344	-0.0234	-0.865	0.0135	29.07	0.0070+
	(0.0239)	(0.0274)	(1.005)	(0.0122)	(17.96)	(0.0041)
Robust P	0.338	0.236	0.247	0.291	0.104	0.137
BW Mean	0.24	0.35	12.23	0.08	91.51	0.02
County N	421	421	421	421	421	421
Individuals	37000	37000	37000	37000	37000	37000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used.

Table A6: Reduced Form Long-Run Effects on Mothers in 1980 Census - 10 Years After Head Start
Without Covariates

	(1) Worked 1975	(2) Employed	(3) Hours	(4) Receive Welfare	(5) Welfare $'s	(6) In School	(7) Years of Ed
A: Single Mothers							
RD_Estimate	-0.0404	-0.0436+	-2.588*	0.0181	131.2+	0.0229**	-0.0333
	(0.0319)	(0.0238)	(1.053)	(0.0219)	(76.73)	(0.0085)	(0.177)
Robust P	0.743	0.363	0.0648	0.216	0.0643	0.0547	0.872
BW Mean	0.50	0.54	19.29	0.36	858.6	0.03	9.94
County N	420	420	420	420	420	420	420
Individuals	11000	11000	11000	11000	11000	11000	11000
B: Non-White Mothers							
RD_Estimate	-0.0559	-0.0912+	-3.727*	-0.0271	57.23	0.0244**	-0.247
	(0.0433)	(0.0496)	(1.602)	(0.0344)	(97.60)	(0.0084)	(0.329)
Robust P	0.920	0.268	0.0444	0.380	0.495	0.0054	0.859
BW Mean	0.48	0.51	17.23	0.28	620.4	0.03	9.80
County N	371	371	371	371	371	371	371
Individuals	17000	17000	17000	17000	17000	17000	17000
C: All Mothers							
RD_Estimate	-0.0140	-0.0138	-0.793	0.0170	76.65	0.0063	-0.0005
	(0.0235)	(0.0224)	(0.847)	(0.0122)	(48.84)	(0.0039)	(0.168)
Robust P	0.577	0.431	0.236	0.363	0.263	0.177	0.851
BW Mean	0.46	0.52	18.80	0.15	358.4	0.02	10.58
County N	420	420	420	420	420	420	420
Individuals	54000	54000	54000	54000	54000	54000	54000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used.

Table A7: Reduced Form Effects on Mother of Head Start Aged Children in 1970 Census Individual Level

	(1) Worked 1965	(2) Employed	(3) Hours	(4) Receive Welfare	(5) Welfare $'s	(6) In School
A: Single Mothers						
RD_Estimate	-0.0150	-0.0893**	-3.829**	-0.0021	7.047	-0.0026
	(0.0353)	(0.0295)	(1.018)	(0.0266)	(43.90)	(0.0092)
Robust P	0.754	0.0155	0.0040	0.857	0.736	0.749
BW Mean	0.36	0.44	15.10	0.37	420.2	0.04
Counties	407	407	407	407	407	407
Individuals (N)	4900	4900	4900	4900	4900	4900
B: Non-White Mothers						
RD_Estimate	-0.0034	-0.0039	-0.371	0.00489	7.098	-0.0003
	(0.0276)	(0.0282)	(1.045)	(0.0153)	(20.80)	(0.0053)
Robust P	0.840	0.979	0.646	0.846	0.679	0.935
BW Mean	0.30	0.40	13.27	0.17	186.7	0.02
Counties	347	347	347	347	347	347
Individuals (N)	12000	12000	12000	12000	12000	12000
C: All Mothers						
RD_Estimate	-0.0052	-0.0166	-0.639	0.0002	6.339	0.0055+
	(0.0161)	(0.0188)	(0.690)	(0.0056)	(8.173)	(0.0029)
Robust P	0.742	0.514	0.457	0.804	0.332	0.0432
BW Mean	0.25	0.35	12.07	0.08	94.65	0.02
Counties	421	421	421	421	421	421
Individuals (N)	37000	37000	37000	37000	37000	37000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used. Covariates are included.

Table A8: Reduced Form Long-Run Effects on Mothers in 1980 Census - 10 Year After Head Start
Individual Level

	(1) Worked 1975	(2) Employed	(3) Hours	(4) Receive Welfare	(5) Welfare $'s	(6) In School	(7) Years of Ed
A: Single Mothers							
RD_Estimate	-0.0274	-0.0124	-1.189	0.0178	134.8+	0.0177**	0.291+
	(0.0214)	(0.0174)	(0.782)	(0.0170)	(70.75)	(0.0052)	(0.156)
Robust P	0.515	0.768	0.221	0.225	0.0883	0.0055	0.142
BW Mean	0.50	0.53	18.84	0.39	919.3	0.03	9.75
Counties	420	420	420	420	420	420	420
Individuals (N)	11000	11000	11000	11000	11000	11000	11000
B: Non-White Mothers							
RD_Estimate	-0.0194	-0.0268	-1.130	0.0257+	123.8*	0.0118*	0.126
	(0.0176)	(0.0188)	(0.791)	(0.0137)	(55.71)	(0.0052)	(0.203)
Robust P	0.428	0.320	0.304	0.0904	0.0574	0.0017	0.584
BW Mean	0.49	0.51	17.40	0.30	694.5	0.029	9.56
Counties	371	371	371	371	371	371	371
Individuals (N)	17000	17000	17000	17000	17000	17000	54000
C: All Mothers							
RD_Estimate	-0.0144	-0.0239*	-1.056*	0.0041	43.71+	0.0062+	0.231+
	(0.0123)	(0.0118)	(0.493)	(0.0062)	(24.84)	(0.0032)	(0.124)
Robust P	0.249	0.0643	0.0422	0.582	0.110	0.107	0.136
BW Mean	0.46	0.52	18.59	0.16	371.7	0.02	10.42
Counties	420	420	420	420	420	420	420
Individuals (N)	54000	54000	54000	54000	54000	54000	54000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Bandwidth is set to 7 percentage points of 1960 county poverty. Regressions are at the county level. Sample consists of all mothers matched to a child born 1964-65 in their household. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). BW mean is the average for the entire sample used. Covariates are included.

Table A9: Continuity of Demographic Characteristics of Mom Sample in 1970

	(1) White	(2) Married	(3) Age	(4) Children Born	(5) Child born Q1	(6) Years of Ed
A: Single Mothers						
RD_Estimate	-0.0343		0.628	0.301	-0.0355	0.359
	(0.0253)		(1.016)	(0.272)	(0.0425)	(0.285)
Robust P	0.148		0.881	0.553	0.825	0.131
BW Mean	0.49		31.69	4.89	0.23	9.21
Counties	407		407	407	407	407
Individuals	4900		4900	4900	4900	4900
B: Non-White Mothers						
RD_Estimate		0.0013	-1.002	-0.334	0.0072	0.363
		(0.0424)	(0.734)	(0.223)	(0.0323)	(0.298)
Robust P		0.382	0.091	0.261	0.386	0.214
BW Mean		0.74	32.22	5.54	0.24	9.40
Counties		347	347	347	347	347
Individuals		12000	12000	12000	12000	12000
C: All Mothers						
RD_Estimate	-0.0387	-0.0029	0.507*	0.183+	0.0012	0.0839
	(0.0481)	(0.0119)	(0.222)	(0.108)	(0.0118)	(0.155)
Robust P	0.559	0.947	0.448	0.287	0.835	0.966
BW Mean	0.72	0.89	31.09	4.67	0.24	10.16
Counties	421	421	421	421	421	421
Individuals	37000	37000	37000	37000	37000	37000

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes all children born between 1964 and the first quarter of 1965. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Bandwidth mean is the mean over the full bandwidth used. Number of counties reported in "N", number of individuals underlying the counties is reported as "Individuals."

Table A10: Continuity of 1980 Demographics

	(1) Population	(2) Black Pop	(3) Pop Aged 5-17	(4) Poverty Rate	(5) Median Income
RD_Estimate	1525.505	862.710	135.650	0.001	125.014
	(5471.135)	(1473.140)	(1430.693)	(0.010)	(352.231)
Bandwidth	7.9	5.7	8.2	4.6	5.6
Robust P-Val	.901	.806	.95	.853	.618
Control Mean	22606	4345	5428	.237	11037
N	471	345	485	277	344

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.001$

Note: Data comes from the 1983 City and County Data Books (which is derived from the Decennial Census).

Table A11: Placebo Composition Tests, No Head Start Aged Children

	(1) Welfare Spell (Mnths)	(2) Spell Under 4mo	(3) Spell Under 12mo	(4) Spell Above Median	(5) Black Black	(6) Mom's Education	(7) Father in Home
A: Optimal Bandwidth							
RD_Estimate	-15.459^{+}	0.045	0.086	-0.091	0.103	-0.574	-0.028
	(8.387)	(0.049)	(0.064)	(0.073)	(0.097)	(0.447)	(0.069)
Robust P-Val	.0508	.414	.186	.181	.296	.214	.73
Bandwidth	5.7	7	5.8	6	6	6.4	6.9
Control Mean	73.9	.0879	.203	.63	.432	7.51	.304
N	281	332	286	292	293	295	329
Individuals	1393	1686	1409	1517	1519	1580	1682
B: Bandwidth = 7							
RD_Estimate	-14.666^{+}	0.045	0.090	-0.090	0.100	-0.572	-0.029
	(7.791)	(0.049)	(0.060)	(0.070)	(0.092)	(0.432)	(0.069)
Robust P-Val	.121	.781	.493	.277	.453	.293	.791
Control Mean	72.4	.0878	.198	.624	.427	7.53	.305
N	336	336	336	336	336	320	336

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes all households with no children aged 3-8 collapsed to the county level. Robust P-Value and optimal bandwidth follow Calonico et al. (2014b). Control mean is the estimate from the local linear regression to the left of the cutoff (the 301st poorest county). Number of counties reported in "N", number of individuals underlying the counties is reported as "Individuals."

Table A12: Birth Month Discontinuity on Kindergarten Attendance: Varying Bandwidth

	(1)	(2)	(3)	(4)	(5)
	All States		K States		
RD_Estimate	0.149**	0.206**	0.209**	0.248**	0.301**
	(0.015)	(0.021)	(0.018)	(0.016)	(0.014)
Robust P-Val	2.34e-08	9.16e-08	5.60e-32	1.67e-16	5.27e-27
Control Mean	.047	.062	.051	.061	.052
Bandwidth	4	4	2	7	10
N	7260	4562	1976	8593	12640

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes AFDC recipient children who reside in states which had universal kindergarten by 1967. The cutoff used is births in or before December 1962. Bandwidth shows how many months of birth are included on either side of the cutoff. Robust P-Value follows Calonico et al. (2014b)

Table A13: Reduced form outcomes of Age RD for Kindergarten on Selection into Welfare

	(1) Spell	(2) Less than 4	(3) Less than 12	(4) Above Median	(5) Moth Edu	(6) Father Home
A: Using Kindergarten Eligibility: States with K programs						
RD_Estimate	-0.864	0.026	0.027	0.001	-0.006	0.008
	(2.518)	(0.021)	(0.032)	(0.034)	(0.173)	(0.028)
Robust P	0.412	0.697	0.579	0.766	0.153	0.757
Control Mean	34.09	0.09	0.32	0.42	9.63	0.22
N	4619	4619	4619	4619	3805	4619
B: Using First Grade Eligibility: All States						
RD_Estimate	-0.636	0.003	-0.055^{+}	-0.004	0.131	-0.036
	(2.516)	(0.020)	(0.032)	(0.034)	(0.169)	(0.028)
Robust P-Val	0.710	0.898	0.450	0.454	0.774	0.736
Control Mean	35.45	0.09	0.35	0.48	9.72	0.25
N	4766	4766	4766	4766	3956	4766
C: Using First Grade Eligibility: Counties near 300 Poorest Cutoff						
RD_Estimate	4.853	0.090	0.077	0.009	-0.580	0.056
	(6.992)	(0.071)	(0.097)	(0.106)	(0.614)	(0.092)
Robust P-Val	0.06	0.65	0.11	0.21	0.43	0.72
Control Mean	29.297	0.070	0.273	0.408	8.061	0.134
N	516	516	516	516	468	516

Note: Conventional standard errors in parenthesis: + p<0.10, * p<0.05, ** p<0.01. Sample includes AFDC recipient mothers who reside in states which had universal kindergarten by 1967. The bandwidth includes mothers whose child was born between September 1962 and April 1963 (bw=4). The cutoff used is births in or before December 1962. Robust P-Value follows Calonico et al. (2014b)